AD708563

Utility Theory for Decision Making

by Peter C. Fishburn

Copy 23 of 130

RAC Research Analysis Corporation

The contents of RAC publications, including the conclusions, represent the views of RAC and should not be considered to have official Department of the Army approval, either expressed or implied, until reviewed and evaluated by that agency and subsequently endorsed.

THIS

PAGE

IS

MISSING

IN

ORIGINAL

DOCUMENT

UTILITY THEORY
FOR
DECISION MAKING

PUBLICATIONS IN OPERATIONS RESEARCH

Operations Research Society of America
Editor for Publications in Operations Research
DAVID B. HERTZ

No. 1. QUEUES, INVENTORIES AND MAINTENANCE
Philip M. Morse

No. 2. FINITE QUEUING TABLES
L. G. Peck and R. N. Hazelwood

No. 3. EFFICIENCY IN GOVERNMENT THROUGH SYSTEMS ANALYSIS
Roland N. McKean

No. 4. A COMPREHENSIVE BIBLIOGRAPHY ON OPERATIONS RESEARCH*
Operations Research Group, Case Institute

No. 5. PROGRESS IN OPERATIONS RESEARCH, VOLUME I
Edited by Russell L. Ackoff

No. 6. STATISTICAL MANAGEMENT OF INVENTORY SYSTEMS*
Harvey M. Wagner

No. 7. PRICE, OUTPUT, AND INVENTORY POLICY*
Edwin S. Mills

No. 8. A COMPREHENSIVE BIBLIOGRAPHY ON OPERATIONS RESEARCH, 1957–1958*
Operations Research Group, Case Institute

No. 9. PROGRESS IN OPERATIONS RESEARCH, VOLUME II
David B. Hertz and Roger T. Eddison

No. 10. DECISION AND VALUE THEORY
Peter C. Fishburn

No. 11. HANDBOOK OF THE POISSON DISTRIBUTION
Frank A. Haight

No. 12. OPERATIONS RESEARCH IN SELLERS' COMPETITION: A STOCHASTIC MICROTHEORY
S. Sankar Sengupta

No. 13. BAYESIAN DECISION PROBLEMS & MARKOV CHAINS
J. J. Martin

No. 14. MATHEMATICAL MODELS OF ARMS CONTROL AND DISARMAMENT: APPLICATION OF MATHEMATICAL STRUCTURES IN POLITICS
Thomas L. Saaty

No. 15. FOURTH INTERNATIONAL CONFERENCE ON OPERATIONS RESEARCH
David B. Hertz and Jacques Melese

No. 16. PROGRESS IN OPERATIONS RESEARCH, VOLUME III
RELATIONSHIP BETWEEN OPERATIONS RESEARCH AND THE COMPUTER
J. S. Aronofsky

No. 17. INTRODUCTION TO SYSTEMS COST EFFECTIVENESS
Karl Seiler, III

No. 18. UTILITY THEORY FOR DECISION MAKING
Peter C. Fishburn

No. 19. THE IMPLEMENTATION OF OPERATIONS RESEARCH
Jan H. B. Huysmans

No. 20. THE CHALLENGE TO SYSTEMS ANALYSIS: PUBLIC POLICY AND SOCIAL CHANGE
Edited by Grace J. Kelleher

No. 21. QUANTITATIVE THEORIES IN ADVERTISING
A. G. Rao

* Out-of-print.

UTILITY THEORY
FOR
DECISION MAKING

PETER C. FISHBURN

Research Analysis Corporation

JOHN WILEY & SONS, INC.

NEW YORK · LONDON · SYDNEY · TORONTO

Library of Congress Catalogue Card Number: 74-107587

SBN 471 26060 6

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

TO MY PARENTS

FOREWORD

This book presents a unified treatment of normative theories for the evaluation of individuals' preferences in a variety of types of decision situations. The material was compiled and developed as part of RAC's Advanced Research Department work program in decision and value theory. Many of the results in the book were developed as a result of basic research investigations under the RAC Institutional Research Program in addition to ONR and ARO support.

NICHOLAS M. SMITH
Head, Advanced Research Department

PREFACE

The underlying motive for this book is the widespread activity of human decision making. Its basic motif is that decisions depend, at least in part, on preferences. Its subject matter is preference structures and numerical representations of preference structures.

Although utility theory has well-recognized roots that extend into the eighteenth and nineteenth centuries, much of its significant growth has occurred in the last two or three decades. This growth, whose major contributions have come out of economics, statistics, mathematics, psychology, and the management sciences, has been greatly stimulated by the use of axiomatic theory. This is evident, for example, in the works of Frank P. Ramsey (1931), John von Neumann and Oskar Morgenstern (1947), Leonard J. Savage (1954), John S. Chipman (1960), and Gerard Debreu (1959, 1960), all of which use the axiomatic approach. In this approach the investigator puts forth a set of axioms or conditions for preferences. It might be said that these conditions characterize a preference structure. Some of them may be viewed as criteria of consistency and coherence for the preferences of a decision maker; others may be viewed as structural and/or simplifying assumptions. In any event, the investigator then seeks to uncover a numerical model that preserves certain characteristics inherent in the assumed preference structure. Further investigation might indicate how such a model can be used to help decision makers examine and perhaps resolve decision problems. This can include methods of estimating the terms (utilities, probabilities) that appear in the model.

During 1963 through 1969, while this book progressed through its own growth and distillation stages, I have been increasingly concerned by the needs for a unifying upper-level text and a research-reference work on utility theory. It is my hope that the book will satisfy these needs for at least the next several years.

The book was written to be self-contained. My experience indicates that many people interested in utility theory are not especially well trained in mathematics. For this reason and to prevent any misunderstanding, I have

included virtually all required background mathematics. This material is introduced when and where it is needed. Those unfamiliar with it will of course find much of it difficult going, but at least I hope they will be spared the trouble of searching elsewhere for it.

Also by way of self-containment, proofs are provided for all but a very few theorems. Browsers will want to skip the proofs, but they are available when desired. In most cases, source credit is given for more involved proofs. In some cases I have expanded others' proofs to make them more accessible to some readers. This is most noticeable with respect to Debreu's additivity theory in Chapter 5 and Savage's expected-utility theory in Chapter 14.

Set theory is the cornerstone mathematics of the text. With no significant exception, all utility theories examined in the book are based on the theory of binary relations. The main binary relation is the preference relation "is preferred to." Algebra, group theory, topology, probability theory, and the theory of mathematical expectation arise at various places.

The exercises are an integral part of the book. Those with boldface numbers cover important material not presented elsewhere in the chapters. Other exercises offer practice on the basic mathematics and on the utility theory and related materials discussed in the chapters. Answers to selected exercises follow Chapter 14. A preview of the book's contents is given in the first chapter.

Finally, you should know about two other books that present a significant amount of material on measurement theory (of which utility theory may be considered a part) that is not found in this book. The first of these is John Pfanzagl's *Theory of Measurement* (John Wiley & Sons, Inc., New York, 1968). The second is being prepared by David H. Krantz, R. Duncan Luce, Patrick Suppes, and Amos Tversky.

McLean, Virginia
June 1969

PETER C. FISHBURN

ACKNOWLEDGMENTS

The preparation of this book was made possible by the joint support of the Department of the Army, DA Contract No. 44-188-ARO-1, and the Office of Naval Research, ONR Contract No. N00014-67-C-0434.

Since 1963 I have been a member of the Advanced Research Department, Research Analysis Corporation. I am extremely grateful to Dr. Nicholas M. Smith, Jr., my department head, and to Frank A. Parker, President of the Research Analysis Corporation, for their continued encouragement and support of my work during this time. I am indebted also to Dr. George S. Pettee, Chairman of RAC's Open Literature Committee, and to his committee, for their guidance. Mrs. Elva Baty typed the first draft of the book in 1963, and it was reviewed by Dr. Irving H. Siegel, Dr. Richard M. Soland, Dr. Jerome Bracken, and Mr. Robert Busacker, all of RAC at that time. The second draft, completed in 1968, was typed by Mrs. Virginia M. Johnson, who also typed the final draft. I offer my sincerest thanks to these friends and many others at RAC who have participated in the preparation of the work.

Dr. David B. Hertz, Editor of this series for ORSA, personally read the material for this book as it was developed. I am grateful to him for his continued encouragement and counsel, and to the Publications Committee of ORSA for its decision to publish.

The original idea for this book came from Professor Russell L. Ackoff. Another great teacher, Professor Leonard J. Savage, helped me to understand his own utility theory and was instrumental in developing the contents of Chapter 10. He made many useful suggestions on the material in Chapters 11 and 13 also. Encouragement and help on the first part of the book came from Professors Gerard Debreu, David H. Krantz, R. Duncan Luce, and Marcel K. Richter.

I am indebted also to many editors and referees. In addition to sources acknowledged in the text, I would like to thank those who have helped me with articles whose contents appear in part in Chapters 9 through 13. A few items in the exercises of Chapter 9 appeared in "Semiorders and Risky

Choices," *Journal of Mathematical Psychology* **5** (1968), 358–361. Chapter 10 grew out of "Bounded Expected Utility," *The Annals of Mathematical Statistics* **38** (1967), 1054–1060. Chapter 11 is based on "Independence in Utility Theory with Whole Product Sets," *Operations Research* **13** (1965), 28–45; "Markovian Dependence in Utility Theory with Whole Product Sets," *Operations Research* **13** (1965), 238–257; "Independence, Trade-Offs, and Transformations in Bivariate Utility Functions," *Management Science* **11** (1965), 792–801; "Stationary Value Mechanisms and Expected Utility Theory," *Journal of Mathematical Psychology* **3** (1966), 434–457; "Conjoint Measurement in Utility Theory with Incomplete Product Sets," *Journal of Mathematical Psychology* **4** (1967), 104–119; "Additive Utilities with Incomplete Product Sets: Application to Priorities and Assignments," *Operations Research* **15** (1967), 537–542; "Interdependence and Additivity in Multivariate, Unidimensional Expected Utility Theory," *International Economic Review* **8** (1967), 335–342; and "A Study of Independence in Multivariate Utility Theory," *Econometrica* **37** (1969), 107–121. Some material in Chapter 12 appeared in "An Abbreviated States of the World Decision Model," *IEEE Transactions on Systems Science and Cybernetics* **4** (1968), 300–306. Chapter 13 is based on "Preference-Based Definitions of Subjective Probability," *Annals of Mathematical Statistics* **38** (1967), 1605–1617, and "A General Theory of Subjective Probabilities and Expected Utilities," *Annals of Mathematical Statistics* **40** (1969), 1419–1429. In addition to these, the material on interval orders in Section 2.4 is based on "Intransitive Indifference with Unequal Indifference Intervals," *Journal of Mathematical Psychology* **7** (1970), and many of the results for strict partial orders throughout the book are summarized in "Intransitive Indifference in Preference Theory: A Survey," *Operations Research* **18** (1970).

Finally and foremost, I thank my wife Janet and our children for their love, and Rebecca and Hummel, who made it possible.

P.C.F.

CONTENTS

UTILITY THEORY
FOR
DECISION MAKING

Chapter 1

INTRODUCTION AND PREVIEW

Decision making serves as the foundation on which utility theory rests. For the purposes of this book we envision a decision maker who must select one alternative (act, course of action, strategy) from a recognized set of decision alternatives. Our study will focus on individuals' preferences in such decision situations. For a connection between decision and preference we shall assume that preferences, to a greater or lesser extent, govern decisions and that, generally speaking, a decision maker would rather implement a more preferred alternative than one that is less preferred.

In the axiomatic systems examined in this book, an individual's preference relation on a set of alternatives enters as a primitive or basic notion. This means that we shall not attempt to define preference in terms of other concepts. We shall, however, suggest that, by self-interrogation, an individual can identify at least some of his preferences.

As we proceed through various types of decision situations it will become apparent that, under specified assumptions, preferences between decision alternatives might be characterized in terms of several factors relating to the alternatives. In cases where alternatives can be viewed as aggregates of several attributes or factors, holistic preferences might be represented as aggregates of preferences on the several factors. In other cases, as in decision under uncertainty, holistic preferences may be represented in terms of utilities for consequences and probabilities for consequences or for "states of the world." These special ways of representing preferences do not of course explain the meaning of the term although they may help in understanding how holistic preferences can be described in terms of other factors.

1.1 GENERAL ORGANIZATION

The three main parts of the text comprise two main divisions of our subject as follows:

Part I. Individual decision under certainty.
Parts II and III. Individual decision under uncertainty.

Part I, titled "Utilities without Probabilities," covers situations where uncertainty is not explicitly formulated. I use the phrase "decision under certainty" as an abbreviation for something like "decision making in which uncertainty, whatever form it might take, is suppressed and not given explicit recognition."

Parts II and III explicitly recognize the form of uncertainty that is characterized by the question: If I implement decision alternative f, then what will happen? Parts II and III differ in their formulations of an uncertain situation, although under appropriate interpretation the two formulations are equivalent. In Part II, titled "Expected-Utility Theory," the uncertainty is expressed in terms of the probability that consequence x will result if act f is implemented. In Part III, "States of the World," uncertainty is expressed in terms of probabilities for contingencies whose occurrence cannot be influenced by the specific act that is implemented but which determine the consequence that results under each available act. The Part II formulation is the one used in Fishburn (1964). The Part III formulation is the one adopted in the version of statistical decision theory sponsored by Savage (1954) and Raiffa and Schlaifer (1961).

In the actual presentations of Parts II and III there is another noticeable difference. In Part III, especially Chapters 13 and 14, the state probabilities as well as the utilities are derived from the preference axioms. In Part II probabilities of acts for consequences are, so to speak, taken as given and enter into the axioms. This is partly rectified in Section 13.4, which presents an axiomatization for the Part II formulation in which the consequence probabilities are derived from the axioms. An alternative axiomatization of the Part II model that also does not use consequence probabilities in the axioms has been developed recently by Duncan Luce and David Krantz. Since this awaits publication as I am completing this book, its important contributions do not appear here.

1.2 PART I: UTILITIES WITHOUT PROBABILITIES

A natural first topic for a study on utility theory is the elementary properties of a preference relation on a set of decision alternatives. The next two chapters go into this in some detail. Their main concern is what might be called the *fundamental theorem of utility*. This has to do with axioms for preferences which guarantee, in a formal mathematical sense, the ability to assign a number (utility) to each alternative so that, for any two alternatives, one is preferred to the other if and only if the utility of the first is greater than the utility of the second.

These two chapters differ primarily in the size assumed for the set of alternatives. Chapter 2 assumes that this set is finite or denumerably infinite;

Chapter 3 covers cases where the alternative set is so large that it is uncountable (neither finite nor denumerable). After dealing with the fundamental theorem, Chapter 2 discusses ordering properties on preferences that are not strong enough to yield the fundamental theorem. Here we shall not assume that indifference ("no preference") is transitive. Along with the fundamental theorem as such, Chapter 3 gives sufficient conditions for order-preserving utilities when the alternative set is a subset of finite-dimensional Euclidean space, and then goes on to consider continuous utility functions.

Additive Utilities

Chapters 4, 5, and 7 deal with cases where each alternative can be viewed as a multiple-factor or multiple-attribute entity. In more mathematical terms, each alternative is an n-tuple of elements, one element from each of a set of n factors. Unlike the other chapters in this trio, Chapter 7 deals explicitly with the case where the n factors are essentially similar. A prototype example for Chapter 7 is the case where n denotes a number of time periods and an alternative specifies income in each period. Time-oriented notions of persistent preferences, impatience, stationarity, and marginal consistency are examined in Chapter 7, as well as a persistent preference difference concept that draws on material in Chapter 6.

Chapters 4 and 5 deal with preference conditions on a set of multiple-factor alternatives that not only yield order-preserving utilities as in Chapters 2 and 3 but also enable the utility of each alternative to be written as the sum of utility numbers assigned to each of the n components of the alternative. In simpler language, these chapters deal with conditions that imply that the utility of a whole can be expressed as the sum of utilities of its parts. In Chapter 4 the alternative set is taken to be finite; in Chapter 5 the number of alternatives is infinite.

Strength of Preference

Chapter 6 is the only chapter in the book that deals primarily with utility concepts involving strength of preference or preference intensity. It is concerned with comparisons between pairs of alternatives and raises the question: Is your difference in preference (degree of preference) between these two alternatives less than, equal to, or greater than your difference in preference between those two alternatives? Chapter 6 is concerned with utility functions that preserve such preference-difference comparisons.

1.3 PARTS II AND III: UTILITIES WITH PROBABILITIES

As noted above, Parts II and III differ in their formulations of decision under uncertainty. Both parts are concerned with simple preference comparisons

between alternatives whose consequences are uncertain, and with preference conditions that not only yield order-preserving utilities for the alternatives but also enable the utility of an alternative to be written as a mathematical expectation involving consequence utilities and consequence probabilities.

In this book, probability is interpreted in a subjective or personal way. Roughly speaking, a probability is a numerical expression of the confidence that a particular person has in the truth of a particular proposition, such as the proposition "if I implement f then consequence x will result," or the proposition "this coin will land 'heads' on the next flip." Such probabilities are required to obey well-defined rules of coherence and consistency. In those cases where probabilities are derived from preference axioms, the primitive notion for probability is preference. Early in Chapter 14 we shall see how probability can be axiomatized in terms of a relation "is less probable than" on a set of propositions or events. Later in Chapter 14 we shall see how "is less probable than" can be defined in terms of "is preferred to." My own viewpoint on probability is heavily influenced by de Finetti (1937) and Savage (1954). Kyburg and Smokler (1964) is recommended for further introductory reading in subjective probability. Chapter 5 of Fishburn (1964) discusses other interpretations of the meaning of probability.

Part II

The first three chapters of Part II derive the expected-utility representation for alternatives with uncertain consequences. In these chapters the consequence probabilities are taken as "givens" so that the alternatives in the preference axioms are probability distributions or measures on a set of consequences. Chapter 8 concentrates on simple probability measures, where each alternative has probability one (certainty) of resulting in a consequence from some finite subset of consequences. Chapter 9 considers simple measures also but, unlike Chapter 8, it does not assume that indifference is transitive. Chapter 10 admits more general probability measures on the consequences.

Uncertainty is combined with multiple-factor consequences in Chapter 11. This chapter identifies conditions that enable the expected utility of an uncertain alternative with n-tuple consequences to be expressed as the sum of expected utilities for each of the n factors. Section 11.4, like Chapter 7, examines the case where the n factors are essentially similar.

Part III

The three chapters in Part III deal with the basic states of the world decision formulation. Chapter 12 introduces this formulation, demonstrates its equivalence to the Part II formulation, and considers some axioms that do not yield the complete expected-utility subjective-probability representation.

Chapter 13, whose material reflects some investigations carried out by Herman Rubin and Herman Chernoff in the late 1940's and early 1950's, presents axioms that yield the complete expected-utility subjective-probability model in the states formulation. Probabilities are used in the axioms of this chapter, but they are extraneous measurement probabilities and not the state probabilities. The latter are derived from the axioms.

Chapter 14 presents Savage's (1954) expected-utility theory. His axioms are free of the extraneous measurement probability device but impose some restrictions on the set of states and the state probabilities that are not imposed by the axioms of Chapter 13.

PART I

UTILITIES WITHOUT PROBABILITIES

With few exceptions, most of the significant developments in individual utility theory for preference structures that do not explicitly incorporate uncertainty or probability have occurred since the beginning of the twentieth century. Economists and mathematical economists are largely, though not exclusively, responsible for these developments. The basic theory (Chapters 2 and 3) deals with the existence of utility functions on a set of alternatives that preserves the ordering of the alternatives based on an individual's preference relation, and with special properties—such as continuity—of utility functions. A secondary basic development (Chapter 6) centers on a strength-of-preference concept that concerns comparisons of preference differences.

Although the assumption of additive utilities for multiple-factor situations (Chapters 4 and 5) was widely used by economists in the mid-nineteenth century, it was discarded by many toward the end of the century. In more recent years, principally since 1959, axiomatic theories for additivity have been developed. These theories show what must be assumed about preferences so that the order-preserving utility functions can be written as combinations of utility functions for the several factors.

Chapter 2

PREFERENCE ORDERS AND UTILITY FUNCTIONS FOR COUNTABLE SETS

Throughout the book we shall let X denote a set whose elements are to be evaluated in terms of preference in a particular decision situation. Depending on the context, the elements in X might be called alternatives, consequences, commodity bundles, cash flows, systems, allocations, inventory policies, strategies, and so forth. This chapter is primarily but not exclusively concerned with cases where X is a *countable* set, which means that X is finite or denumerable. A set is *denumerable* if and only if its elements can be placed in one-to-one correspondence with the elements in the set $\{1, 2, 3, \ldots\}$ of positive integers. The set $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$ of all integers and the set of rational numbers (expressible as ratios of integers) are denumerable.

Throughout the book we shall take strict preference $\prec$ as the basic binary relation on X or on a set based on X, and indifference $\sim$ will be defined as the absence of strict preference. One could also begin with a preference-indifference relation $\preccurlyeq$ (read $x \preccurlyeq y$ as x is not preferred to y), but I have come to prefer $\prec$ for several technical reasons plus the fact that we tend to think in terms of preference rather than preference-indifference.

The first main result of this chapter is that, when X is countable, numbers $u(x), u(y), \ldots$ can be assigned to the elements $x, y, \ldots$ in X in such a way that

$$x \prec y \Leftrightarrow u(x) < u(y)$$

holds if $\prec$ on X is a weak order (Definition 2.1). The $\Leftrightarrow$ means "if and only if" and its companion $\Rightarrow$ means "implies." A second main result says that there is a real-valued function u on X such that

$$x \prec y \Rightarrow u(x) < u(y)$$

when $\prec$ on X is a strict partial order (Definition 2.2), provided that X is countable. Several other utility-representation theorems are presented later in the chapter.

2.1 BINARY RELATIONS

The entire book is based on binary relations. A *binary relation* on a set Y is a set of ordered pairs (x, y) with $x \in Y$ and $y \in Y$. The $x \in Y$ means that x is an element in Y; we often abbreviate $x \in Y, y \in Y$ by writing $x, y \in Y$.

The universal binary relation on Y is the set $\{(x, y): x, y \in Y\}$ of all ordered pairs from Y. In general $\{x: S\}$ is the set of all elements x that satisfy the conditions specified by S. If R is a binary relation on Y then it is a subset of the universal binary relation. In general, $A \subseteq B$ (A is a subset of B) means that every element in A is in B also.

We often write xRy to mean that $(x, y) \in R$. Similarly, not xRy (it is false that x stands in the relation R to y) means that $(x, y) \notin R$. In general $a \notin A$ means that a is not an element in A. If R is a binary relation on Y then for each (x, y) in the universal relation either xRy or not xRy, and not both.

Because we are dealing with *ordered* pairs, (x, y) is not the same as (y, x) unless $x = y$. Hence, if R is a binary relation on Y and if $x, y \in Y$, then exactly one of the following four cases holds:

1. (xRy, yRx),
2. $(xRy$, not $yRx)$,
3. (not $xRy, yRx)$,
4. (not xRy, not $yRx)$.

Let Y be the set of all living people. Let R_1 mean "is shorter than," so that xR_1y means that x is shorter than y. Case (1) is impossible. Case (2) holds when x is shorter than y. Case (4) holds when x and y are of equal height. R_1 is an example of a weak order.

Next, let R_2 be "is the brother of" (by having at least one parent in common). Here cases (2) and (3) are impossible. R_2 is not transitive since if xR_2y and yR_2z it does not necessarily follow that xR_2z. (Why?)

Some Relation Properties

The binary relations we use will be assumed to have certain properties. A list of some of these follows. A binary relation R on a set Y is

$p1$. *reflexive* if xRx for every $x \in Y$,
$p2$. *irreflexive* if not xRx for every $x \in Y$,
$p3$. *symmetric* if $xRy \Rightarrow yRx$, for every $x, y \in Y$,
$p4$. *asymmetric* if $xRy \Rightarrow$ not yRx, for every $x, y \in Y$,
$p5$. *antisymmetric* if $(xRy, yRx) \Rightarrow x = y$, for every $x, y \in Y$,
$p6$. *transitive* if $(xRy, yRz) \Rightarrow xRz$, for every $x, y, z \in Y$,
$p7$. *negatively transitive* if (not xRy, not $yRz) \Rightarrow$ not xRz, for every $x, y, z \in Y$,

p8. connected or *complete* if xRy or yRx (possibly both) for every $x, y \in Y$,
p9. weakly connected if $x \neq y \Rightarrow (xRy$ or $yRx)$ throughout Y.

Several other properties are introduced in Section 2.4.

An asymmetric binary relation is irreflexive. An irreflexive and transitive binary relation is asymmetric: if (xRy, yRx) then $p6$ gives xRx, which violates $p2$. It is also useful to note that R is negatively transitive if and only if, for all $x, y, z \in Y$,

$$xRy \Rightarrow (xRz \text{ or } zRy). \tag{2.1}$$

To prove this suppose first that, in violation of (2.1), $(xRy$, not xRz, not $zRy)$. Then, if the $p7$ condition holds, we get not xRy, which contradicts xRy. Hence the $p7$ condition implies (2.1). On the other hand, suppose the $p7$ condition fails with (not xRy, not yRz, xRz). Then (2.1) must be false. Hence (2.1) implies the $p7$ condition.

The relation R_1 (shorter than) is irreflexive, asymmetric, transitive, and negatively transitive. If no two people are of equal height, R_1 is weakly connected. R_2 (brother of) is symmetric.

2.2 PREFERENCE AS A WEAK ORDER

Binary relations that have or are assumed to have certain properties are often given special names. In this section we shall be most concerned with three types of binary relations, namely weak orders, strict orders, and equivalences.

Definition 2.1. A binary relation R on a set Y is

a. a *weak order* $\Leftrightarrow R$ on Y is asymmetric and negatively transitive;
b. a *strict order* $\Leftrightarrow R$ on Y is a weakly connected weak order;
c. an *equivalence* $\Leftrightarrow R$ on Y is reflexive, symmetric, and transitive.

The relation $<$ on the real numbers is a weak order and also a strict order since $x < y$ or $y < x$ whenever $x \neq y$; $=$ on the real numbers is an equivalence, since $x = x$, $x = y \Rightarrow y = x$, and $(x = y, y = z) \Rightarrow x = z$.

An equivalence on a set defines a natural partition of the set into a class of disjoint, nonempty subsets, such that two elements of the original set are in the same class if and only if they are equivalent. These classes are called *equivalence classes*. Let

$$R(x) = \{y : y \in Y \text{ and } yRx\}.$$

If R is an equivalence then $R(x)$ is the *equivalence class generated by* x. In this

case you can readily show that $R(x) = R(y)$ if and only if xRy. Thus, any two equivalence classes are either identical or disjoint (have no element in common). When R on Y is an equivalence, we shall denote the set of equivalence classes of Y under R as Y/R.

Preference as a Weak Order

Taking preference $<$ as basic (read $x < y$ as x is less preferred than y, or y is preferred to x) we shall define indifference $\sim$ as the absence of strict preference:

$$x \sim y \Leftrightarrow (\text{not } x < y, \text{ not } y < x). \tag{2.2}$$

Indifference might arise in several ways. First, an individual might truly feel that, in a preference sense, there is no real difference between x and y. He would just as soon have x as y and vice versa. Secondly, indifference could arise when the individual is uncertain as to his preference between x and y. He might find the comparison difficult and may decline to commit himself to a strict preference judgment while not being sure that he regards x and y as equally desirable (or undesirable). Thirdly, $x \sim y$ might arise in a case where the individual considers x and y incomparable (in some sense) on a preference basis.

Asymmetry is an "obvious" condition for preference. It can be viewed as a criterion of consistency. If you prefer x to y, you should not simultaneously prefer y to x.

Transitivity is implied by asymmetry and negative transitivity, and it seems like a reasonable criterion of coherence for an individual's preferences. If you prefer x to y and prefer y to z, common sense suggests that you should prefer x to z.

However, the full force of weak order is open to criticism since it imparts a rather uncanny power of preferential judgment to the individual, as can be seen from (2.1). To see how (2.1) might fail, suppose that in a funding situation you feel that \$1000 is about the best allocation. Your preference decreases as you move away from \$1000 in either direction. Although you prefer \$955 to \$950, it may also be true that you have no sure preference between \$950 and \$1080 or between \$955 and \$1080. Then (\$950 $<$ \$955, \$950 $\sim$ \$1080, \$955 $\sim$ \$1080) in violation of (2.1).

In this example, indifference is not transitive. Armstrong (1950, p. 122) speaks of intransitive indifference as arising from "the imperfect powers of discrimination of the human mind whereby inequalities become recognizable only when of sufficient magnitude." Later sections of this chapter take account of such limited discriminatory powers by not requiring $\sim$ to be transitive.

Our first theorem notes several consequences of weak order, including the

transitivity of indifference. For this theorem and for later work we shall define preference-indifference $\precsim$ as the union of $\prec$ and $\sim$:

$$x \precsim y \Leftrightarrow x \prec y \quad \text{or} \quad x \sim y. \tag{2.3}$$

THEOREM 2.1. *Suppose $\prec$ on X is a weak order, being asymmetric and negatively transitive. Then*

a. exactly one of $x \prec y$, $y \prec x$, $x \sim y$ holds for each $x, y \in X$;
b. $\prec$ is transitive;
c. $\sim$ is an equivalence (reflexive, symmetric, transitive);
d. $(x \prec y, y \sim z) \Rightarrow x \prec z$, and $(x \sim y, y \prec z) \Rightarrow x \prec z$;
e. $\precsim$ is transitive and connected;
f. with $\prec'$ on $X/\sim$ (the set of equivalence classes of X under $\sim$) defined by

$$a \prec' b \Leftrightarrow x \prec y \quad \textit{for some} \quad x \in a \quad \textit{and} \quad y \in b, \tag{2.4}$$

$\prec'$ on $X/\sim$ is a strict order.

Proof. Part (a) follows from asymmetry and (2.2). For (b), suppose $x \prec y$ and $y \prec z$. Then, by (2.1), ($x \prec z$ or $z \prec y$) and ($y \prec x$ or $x \prec z$). Since $z \prec y$ and $y \prec x$ are false by asymmetry, $x \prec z$. Thus $\prec$ is transitive. Suppose $x \sim y$, $y \sim z$, and not $x \sim z$, in violation of the transitivity of $\sim$. Then, by (a), either $x \prec z$ or $z \prec x$, so that by (2.1) one of $x \prec y$, $y \prec z$, $z \prec y$, and $y \prec x$ must hold, which contradicts $x \sim y$, $y \sim z$, and (a). Hence $\sim$ is transitive. Suppose as in (d) that $x \prec y$ and $y \sim z$. Then, by (a) and (2.1), $x \prec z$. The second half of (d) is similarly proved. For (e) the transitivity of $\precsim$ follows immediately from (b), (c), and (d). For the completeness of $\precsim$ suppose to the contrary that (not $x \precsim y$, not $y \precsim x$). Then, by (2.3), (not $x \prec y$, not $x \sim y$, not $y \prec x$), which violates (a).

Finally, we examine the properties of a strict order for $\prec'$ on $X/\sim$:

1. asymmetry. If $a \prec' b$ and $b \prec' a$ then $x \prec y$ and $y' \prec x'$ for some $x, x' \in a$ and $y, y' \in b$, with $x \sim x'$ and $y \sim y'$. By (d), $x' \prec y$. Again, by (d), $x' \prec y'$, which contradicts $y' \prec x'$.

2. negative transitivity. Suppose $a \prec' b$ with $x \in a$, $y \in b$, and $x \prec y$. For any $c \in X/\sim$ and any $z \in c$, (2.1) implies that $x \prec z$ (in which case $a \prec' c$) or that $z \prec y$ (in which case $c \prec' b$).

3. weak connectedness. Suppose $a, b \in X/\sim$ and $a \neq b$. Then a and b are disjoint so that if $x \in a$ and $y \in b$ then not $x \sim y$. Hence, by (a) either $x \prec y$ or $y \prec x$, so that either $a \prec' b$ or $b \prec' a$. ◆

An Order-Preserving Utility Function

THEOREM 2.2. *If $<$ on X is a weak order and $X/\sim$ is countable then there is a real-valued function u on X such that*

$$x < y \Leftrightarrow u(x) < u(y), \qquad \textit{for all } x, y \in X. \tag{2.5}$$

The utility function u in (2.5) is said to be order-preserving since the numbers $u(x)$, $u(y)$, ... as ordered by $<$ faithfully reflect the order of $x, y, \ldots$ under $\prec$. Clearly, if (2.5) holds, then

$$x \prec y \Leftrightarrow v(x) < v(y), \qquad \text{for all } x, y \in X,$$

for a real-valued function v on X if and only if $[v(x) < v(y) \Leftrightarrow u(x) < u(y)]$ holds throughout X. In the next section we shall consider the case where $\Leftrightarrow$ in (2.5) must be replaced by $\Rightarrow$. In later chapters we shall meet utility functions with properties beyond that of order preservation.

Under the conditions of Theorem 2.2, (2.5) implies that, for all $x, y \in X$, $x \sim y \Leftrightarrow u(x) = u(y)$, and $x \preccurlyeq y \Leftrightarrow u(x) \leq u(y)$, where $\sim$ and $\preccurlyeq$ are defined by (2.2) and (2.3) respectively.

The following proof of the theorem is similar to proofs given by Birkhoff (1948, p. 31) and Suppes and Zinnes (1963, pp. 26–28). As we shall see in Chapter 3, the conclusion of the theorem can be false when $X/\sim$ is uncountable (neither finite nor denumerable).

Proof of Theorem 2.2. Assuming the hypotheses of the theorem we shall assume also that $X/\sim$ is denumerable. The $X/\sim$ finite proof is similar and is left to the reader. Let the elements in $X/\sim$ be enumerated as $a_1, a_2, a_3, \ldots$ and let the rational numbers be enumerated as $r_1, r_2, r_3, \ldots$. No particular $\prec'$ ordering (see (2.4)) or $<$ ordering is implied by these enumerations. We define a real-valued function u on $X/\sim$ as follows, recalling that $\prec'$ as in (2.4) is a strict order on $X/\sim$.

Set $u(a_1) = 0$. For a_m it follows from the properties for $\prec'$ and induction that exactly one of the following holds:

1. $a_i \prec' a_m$ for all $i < m$: if so, set $u(a_m) = m$,
2. $a_m \prec' a_i$ for all $i < m$: if so, set $u(a_m) = -m$,
3. $a_i \prec' a_m \prec' a_j$ for some $i, j < m$ and not $(a_i \prec' a_h \prec' a_j)$,

for every positive integer h that is less than m and differs from i and j: if so, set $u(a_m)$ equal to the first r_k in the enumeration $r_1, r_2, r_3, \ldots$ for which $u(a_i) < r_k < u(a_j)$. Such an r_k exists since there is a rational number between any two different numbers.

By construction, $u(a_m) \neq u(a_i)$ for all $i < m$, and $a_i \prec' a_j \Leftrightarrow u(a_i) < u(a_j)$ for all $i, j \leq m$. This holds for every positive integer m. Hence it holds on all

of $X/\sim$. Finally, define u on X by

$$u(x) = u(a) \quad \text{whenever} \quad x \in a.$$

Equation (2.5) then follows provided that, when $a <' b$, $x < y$ for every $x \in a$ and $y \in b$, which follows directly from (2.4) and Theorem 2.1(d). ◆

As you will easily note, if (2.5) holds, then $<$ on X must be a weak order. Hence if $<$ on X is not a weak order then (2.5) is impossible regardless of the size of X.

2.3 PREFERENCE AS A STRICT PARTIAL ORDER

Throughout the rest of this chapter we shall look at cases where indifference is not assumed to be transitive. This section considers the case where $<$ is a strict partial order.

Definition 2.2. A binary relation R on a set Y is a *strict partial order* if and only if it is irreflexive and transitive.

Since this allows $(x \sim y, y \sim z, x < z)$ when $<$ on X is a strict partial order, $\sim$ is not necessarily transitive and therefore may not be an equivalence. However, a new relation $\approx$, defined as

$$x \approx y \Leftrightarrow (x \sim z \Leftrightarrow y \sim z, \text{ for all } z \in X) \tag{2.6}$$

does turn out to be transitive when $<$ is a strict partial order. $x \approx y$ holds if, whenever x is indifferent to a $z \in X$, y also is indifferent to z, and vice versa. For comparison with Theorem 2.1 we have the following.

THEOREM 2.3. *Suppose $<$ on X is a strict partial order, being irreflexive and transitive. Then*

a. exactly one of $x < y, y < x, x \approx y$, $(x \sim y$, not $x \approx y)$ holds for each $x, y \in X$;
b. $\approx$ is an equivalence;
c. $x \approx y \Leftrightarrow (x < z \Leftrightarrow y < z$ and $z < x \Leftrightarrow z < y$, for all $z \in X)$;
d. $(x < y, y \approx z) \Rightarrow x < z$, and $(x \approx y, y < z) \Rightarrow x < z$;
e. with $<^*$ on $X/\approx$ (the set of equivalence classes of X under $\approx$) defined by

$$a <^* b \Leftrightarrow x < y \text{ for some } x \in a \text{ and } y \in b, \tag{2.7}$$

$<^*$ on $X/\approx$ is a strict partial order.

Proof. (a) follows from asymmetry (implied by irreflexivity and transitivity) and the fact that $x \approx y$ can hold only if $x \sim y$. For (b), the reflexivity

and symmetry of $\approx$ follow directly from (2.6) and the reflexivity and symmetry of $\sim$. Suppose $x \approx y$ and $y \approx z$. Then, by (2.6), if $x \sim t$ then $y \sim t$ and, again by (2.6), if $y \sim t$ then $z \sim t$. Hence $x \sim t \Rightarrow z \sim t$. Conversely $z \sim t \Rightarrow x \sim t$, so that $x \approx z$ as desired for transitivity.

For part (*c*) suppose first that $x \approx y$. If $x < z$ then either $y < z$ or $y \sim z$, for if $z < y$ then $x < y$ by transitivity of $<$. But if $y \sim z$ then $x \sim z$ by (2.6), which contradicts $x < z$. Hence $x < z \Rightarrow y < z$. Similarly $y < z \Rightarrow x < z$. A similar proof shows that $z < x \Leftrightarrow z < y$. (This also establishes (*d*).) On the other hand, assume that the right part of (*c*) holds. Then, if $x \sim t$, it cannot be true that either $y < t$ or $t < y$ so that $y \sim t$ by (2.1) and the asymmetry of $<$. Conversely $y \sim t \Rightarrow x \sim t$. Hence $x \approx y$.

For (*e*), we cannot have $a <^* a$ when $a \in X/\approx$, for then $x < y$ for some x and y for which $x \approx y$, which is false by (*a*). For transitivity suppose $(a <^* b, b <^* c)$. Then $(x < y, y \approx y', y' < z)$ for some $x \in a, y, y' \in b$, and $z \in c$. $x < z$ then follows from (*d*) so that $a <^* c$. ◆

Zorn's Lemma and Szpilrajn's Extension Theorem

Before we can establish a utility-representation theorem for the case where $<$ is a strict partial order and $X/\approx$ is countable, we need to prove the following theorem, due to Szpilrajn (1930).

THEOREM 2.4. *If $<^*$ is a strict partial order on a set Y then there is a strict order $<^0$ on Y that includes $<^*$, so that*

$$x <^* y \Rightarrow x <^0 y, \qquad \textit{for all } x, y \in Y. \tag{2.8}$$

The utility theorem given later as Theorem 2.5 is very easily proved from Theorem 2.4 and the proof of Theorem 2.2.

To establish Szpilrajn's theorem, which holds regardless of the size of Y, we shall need an axiom of set theory that goes by the name of Zorn's Lemma.

ZORN'S LEMMA. *Suppose P on Y is a strict partial order and, for any subset Z of Y on which P is a strict order, there is a $y \in Y$ such that zPy or $z = y$ for all $z \in Z$. Then there is a $y^* \in Y$ such that y^*Px for no $x \in Y$.*

Consider the real numbers in their natural order under $<$. Since $<$ itself on the numbers is a strict order but there is no number y such that $x < y$ or $x = y$ for every number x, the "lemma" does not imply that the real numbers have a maximal element under $<$, as of course they do not.

Zorn's Lemma, used today by most mathematicians, is an assumption. Kelley (1955, pp. 31–36) presents other axioms that are equivalent to Zorn's

Lemma. One of these is the Axiom of Choice: if $\mathcal{S}$ is a set of nonempty sets then there is a function f on $\mathcal{S}$ such that $f(S) \in S$ for each $S \in \mathcal{S}$.

Proof of Theorem 2.4. If $<^*$ is a strict order, there is nothing to prove. Suppose then that $<^*$ is a strict partial order and that x, y in Y are such that $x \not\sim^* y$, (not $x <^* y$, not $y <^* x$). Define $<^1$ on Y thus:

$$a <^1 b \Leftrightarrow a <^* b \quad \text{or else} \quad [(a <^* x \text{ or } a = x), (y <^* b \text{ or } y = b)]. \tag{2.9}$$

Clearly, $a <^* b \Rightarrow a <^1 b$, and $x <^1 y$. We prove first that $<^1$ is a strict partial order.

A. $<^1$ is irreflexive. To the contrary suppose $a <^1 a$. Then, if either $(a <^* x, y <^* a)$ or $(a <^* x, y = a)$ or $(a = x, y <^* a)$, we get $y <^* x$, which is false. Also, $a <^* a$ and $(a = x, y = a)$ are false by assumption. Hence $a <^* a$ is false.

B. $<^1$ is transitive. Assume $(a <^1 b, b <^1 c)$. If $(a <^* b, b <^* c)$ then $a <^* c$ so that $a <^1 c$. If $(a <^* b, (b <^* x \text{ or } b = x)$ and $(y <^* c \text{ or } y = c))$ then $a <^* x$ so that $a <^1 c$ by (2.9). If $((a <^* x \text{ or } a = x)$ and $(y <^* b$ or $y = b), b <^* c)$ then $y <^* c$ and hence $a <^1 c$. Finally, if neither $a <^* b$ nor $b <^* c$ then, by (2.9), $(y <^* b \text{ or } y = b)$ from $a <^1 b$ and $(b <^* x \text{ or } b = x)$ from $b <^1 c$, which are incompatible since they give $y <^* x$ or $y = x$, which are false. Hence this final case cannot arise.

We now use Zorn's Lemma. With $A \subseteq B \Leftrightarrow A$ is a subset of B, we define $A \subset B \Leftrightarrow (A \subseteq B, \text{ not } B \subseteq A)$. Let $\mathfrak{R}$ be the set of all strict partial orders on Y that include $<^*$, so that $R \in \mathfrak{R} \Leftrightarrow (R$ on Y is a strict partial order and $<^* \subseteq R)$. In Zorn's Lemma as stated above, $\subset$ takes the part of P and $\mathfrak{R}$ takes the part of Y.

Clearly, $\subset$ on $\mathfrak{R}$ is a strict partial order. Let $\mathcal{S}$ be a subset of $\mathfrak{R}$ on which $\subset$ is a strict order. (We omit the trivial case where $\mathcal{S} = \varnothing$.) Let S be the set of all (x, y) that are in at least one $R \in \mathcal{S}$: that is, $(x, y) \in S$ or xSy if and only if $(x, y) \in R$ or xRy for some $R \in \mathcal{S}$. Clearly, $R \subseteq S$ for every $R \in \mathcal{S}$. To apply Zorn's Lemma we need to show that $S \in \mathfrak{R}$, or that S on Y is a strict partial order:

A. S is irreflexive. $(x, x) \notin S$ since $(x, x) \notin R$ for every $R \in \mathfrak{R}$.

B. S is transitive. If $(x, y) \in S$ and $(y, z) \in S$ then $(x, y) \in S_1$ and $(y, z) \in S_2$ for some S_1 and S_2 in $\mathcal{S}$. For definiteness suppose $S_1 \subseteq S_2$. Then $(x, y) \in S_2$ and hence $(x, z) \in S_2$ by transitivity, so that $(x, z) \in S$ by the definition of S.

It follows from Zorn's Lemma that there is a $<^0 \in \mathfrak{R}$ such that $<^0 \subset R$ for no $R \in \mathfrak{R}$. Because $<^0$ is in $\mathfrak{R}$, it is a strict partial order. To show that it is a strict order, it remains to note that $<^0$ on Y is weakly connected, for when this is true $<^0$ must be a strict order. (You can easily show that a

weakly connected strict partial order satisfies (2.1), or negative transitivity, and is thus a strict order by Definition 2.1.) Suppose then that contrary to weak connectedness there are $x, y \in Y$ with $x \neq y$ and (not $x <^0 y$, not $y <^0 x$). Then, by the first part of this proof, there is a strict partial order $<^1$ on Y such that $a <^0 b \Rightarrow a <^1 b$, and $x <^1 y$. But then $<^0 \subset <^1$ which contradicts $<^0 \subset R$ for no $R \in \mathcal{R}$. Hence $<^0$ is weakly connected. ◆

Another Utility Theorem

With $\approx$ defined by (2.6), the following theorem says that when $<$ is irreflexive and transitive and $X/\approx$ is countable, numbers can be assigned to the elements of X so as to faithfully preserve both $<$ and $\approx$. However, because $\sim$ can be intransitive, we cannot guarantee that $u(x) = u(y)$ when $x \sim y$ and not $x \approx y$. We might have any one of $u(x) = u(y)$, $u(x) < u(y)$, and $u(y) < u(x)$ when ($x \sim y$, not $x \approx y$).

THEOREM 2.5. *If $<$ on X is a strict partial order and $X/\approx$ is countable then there is a real-valued function u on X such that, for all $x, y \in X$,*

$$x < y \Rightarrow u(x) < u(y) \tag{2.10}$$

$$x \approx y \Rightarrow u(x) = u(y). \tag{2.11}$$

Proof. By Theorem 2.3(*e*), $<^*$ on $X/\approx$ as defined in (2.7) is a strict partial order. By Theorem 2.4, there is a strict order $<^0$ on $X/\approx$ that includes $<^*$. With $X/\approx$ countable, the proof of Theorem 2.2 guarantees a real-valued function u on $X/\approx$ such that $a <^0 b \Leftrightarrow u(a) < u(b)$, for all $a, b \in X/\approx$. With $a \in X/\approx$, set $u(x) = u(a)$ whenever $x \in a$. Then, if $x \approx y$, $u(x) = u(y)$, so that (2.11) holds. And if $x < y$ with $x \in a$ and $y \in b$ then $a <^* b$ by (2.7) and Theorem 2.3(*d*): hence $a <^0 b$ so that $u(a) < u(b)$ and $u(x) < u(y)$. ◆

2.4 ORDERED INDIFFERENCE INTERVALS

There are other interesting assumptions for preferences that add things to strict partial order, but still retain the possibility of intransitive indifference. Two such conditions were introduced into preference theory by Luce (1956). They are stated here in the form given by Scott and Suppes (1958, p. 117).

*p*10. $(x < y, z < w) \Rightarrow (x < w \text{ or } z < y)$, for all $x, y, z, w \in X$.
*p*11. $(x < y, y < z) \Rightarrow (x < w \text{ or } w < z)$, for all $x, y, z, w \in X$.

It is easily seen that if $<$ is irreflexive and either *p*10 or *p*11 holds then $<$ is transitive. When $<$ is a strict partial order, the only instances of *p*10 and *p*11 that are not already implied by irreflexivity and transitivity are those illustrated in Figure 2.1. For *p*10, we have the case shown on the left of the figure

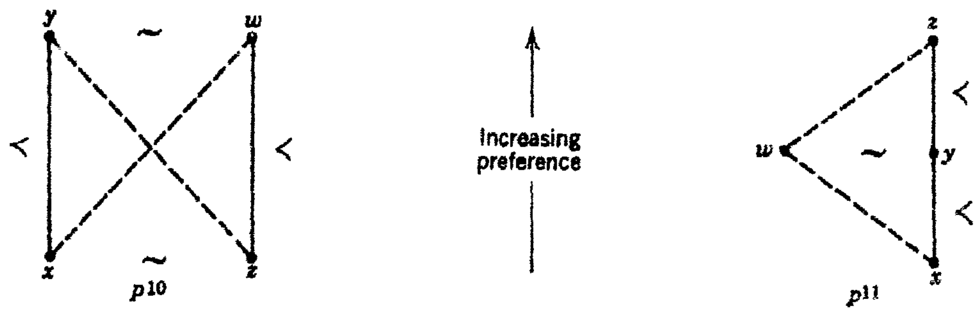

Figure 2.1 Cases not covered by irreflexivity and transitivity.

where $x \prec y$, $z \prec w$, $x \sim z$, and $y \sim w$, with $x \neq z$ and $y \neq w$. In a sort of cross-connectedness, $p10$ says that at least one of the dashed lines must be strict preference: we can't have both $x \sim w$ and $y \sim z$. For $p11$ we get the picture on the right of the figure where $x \prec y \prec z$ and $w \sim y$ with $w \neq y$. Here $p11$ says that at least one of the dashed w-lines must represent strict preference: w can't be indifferent to each of x, y, and z.

Conditions $p10$ and $p11$ may seem reasonable if the elements of X are naturally ordered and preference is either nondecreasing or nonincreasing as one proceeds along the natural order. For example, if you prefer your coffee black it seems fair to assume that your preference will not increase as x, the number of grains of sugar in your coffee, increases. You might well be indifferent between $x = 0$ and $x = 1$, between $x = 1$ and $x = 2, \ldots$, but of course will prefer $x = 0$ to $x = 1000$. Although $\sim$ is not transitive here, $p10$ and $p11$ would probably hold along with irreflexivity.

However, if there are several factors that influence preference or if there is only one basic factor along which preference increases up to a point and decreases thereafter, $\prec$ may fail to satisfy the cases of $p10$ and $p11$ shown in Figure 2.1. To continue with coffee and sugar, suppose you like about 1000 grains of sugar in your coffee. The left part of Figure 2.2 shows a case where it might be true that $x \prec y$, $z \prec w$, $x \sim w$, and $y \sim z$, in violation of $p10$. The right part of the figure suggests that $p11$ may fail with $x \prec y \prec z$ and

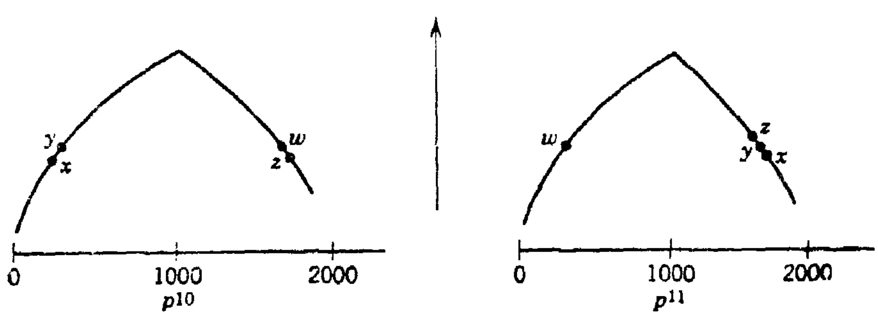

Figure 2.2 "Failures" of $p10$ and $p11$ for single-peaked preferences.

$w \sim x$, $w \sim y$, $w \sim z$. We could expect both $p10$ and $p11$ to hold on a fixed side of your peak or ideal (Coombs, 1964) but there seems to be little reason to suppose that they hold for the cases illustrated. The funding situation in Section 2.2 gives another peaked situation where $p10$ and $p11$ might not hold.

Definition 2.3. A binary relation is an *interval order* if it is irreflexive and satisfies $p10$, and a *semiorder* if it is irreflexive and satisfies $p10$ and $p11$.

The term "semiorder" was introduced by Luce (1956) and is now standard terminology. The way I use "interval order" is not standard, but seems reasonable in view of Theorem 2.7.

Interval Orders

In the rest of the chapter, $\approx$ is defined by (2.6). For interval orders ($p2$, $p10$) we shall use the following:

$$x <^1 y \Leftrightarrow (x \sim z, z < y) \quad \text{for some } z \in X \tag{2.12}$$

$$x <^2 y \Leftrightarrow (x < z, z \sim y) \quad \text{for some } z \in X. \tag{2.13}$$

THEOREM 2.6. *If $<$ on X is an interval order then each of $<^1$ and $<^2$ is a weak order, and $x \approx y \Leftrightarrow (x =^1 y, x =^2 y)$, where $x =^i y \Leftrightarrow$ (not $x <^i y$, not $y <^i x$).*

Proof. The final assertion follows from (2.6). To prove asymmetry for $<^1$ suppose to the contrary that $(x <^1 y, y <^1 x)$. Then $(x \sim z, z < y)$ and $(y \sim w, w < x)$ for some $z, w \in X$, which contradict $p10$. To establish negative transitivity suppose to the contrary that (not $x <^1 y$, not $y <^1 z$, $x <^1 z$). By $x <^1 z$, $(x \sim t, t < z)$ for some $t \in X$. From $x \sim t$ and not $x <^1 y$, (2.12) implies not $t < y$. From $t < z$ and not $y <^1 z$, (2.12) yields not $t \sim y$. Hence $y < t$. But then, by transitivity, $y < z$ which implies $y <^1 z$, contradicting not $y <^1 z$. Hence $<^1$ is negatively transitive. The proof for $<^2$ is similar and is left to the reader. ◆

THEOREM 2.7. *If $<$ on X is an interval order and $X/\approx$ is countable then there are real-valued functions u and σ on X with $\sigma(x) > 0$ for all $x \in X$ such that*

$$x < y \Leftrightarrow u(x) + \sigma(x) < u(y), \quad \textit{for all } x, y \in X. \tag{2.14}$$

Note also that if (2.14) holds then $p10$ must hold.

Theorem 2.7 is like the weak order Theorem 2.2 with the addition of a "vagueness" function σ which allows for intransitive indifference. The indifference interval for x is $I(x) = [u(x), u(x) + \sigma(x)]$. By (2.14), $I(x)$ is

wholly to the left of $I(y)$ if and only if $x < y$. If two intervals intersect then their elements are indifferent. As seen by the failure $(x < y < z, w \sim x, w \sim y, w \sim z)$ of $p11$, one indifference interval may lie entirely within another interval: in the case at hand, $I(y)$ must be shorter than $I(w)$.

Proof of Theorem 2.7. Let $<$ on X be an interval order. Using the Axiom of Choice let Y consist of one element from each equivalence class in $X/\approx$. For each $x \in Y$ let x^* denote an artificial element that corresponds to x, with Y^* the set of artificial elements. Define $<^3$ on $Y \cup Y^*$ (the set of elements in Y or Y^*) as follows:

$$x <^3 y \Leftrightarrow x <^1 y \tag{2.15}$$

$$x^* <^3 y^* \Leftrightarrow x <^2 y \tag{2.16}$$

$$x^* <^3 y \Leftrightarrow x < y \tag{2.17}$$

$$x <^3 y^* \Leftrightarrow x \preccurlyeq y \tag{2.18}$$

where $\preccurlyeq \; = \; < \cup \sim$ as in (2.3). We prove that $<^3$ on $Y \cup Y^*$ is a weak order.

Asymmetry. We want $a <^3 b \Rightarrow$ not $b <^3 a$. If $(a, b) = (x, y)$ or $(a, b) = (x^*, y^*)$ then asymmetry follows from Theorem 2.6 and (2.15) or (2.16). Suppose $(a, b) = (x^*, y)$ and $(a <^3 b, b <^3 a)$. Then $(x < y, y \preccurlyeq x)$ by (2.17) and (2.18), which is impossible.

Negative Transitivity. We shall suppose that (not $a <^3 b$, not $b <^3 c$, $a <^3 c$) and obtain a contradiction. The cases for $(a, b, c) = (x, y, z)$ and $(a, b, c) = (x^*, y^*, z^*)$ are covered by Theorem 2.6. The others follow.

1. (x, y, z^*). Then not $x <^1 y$, $z < y$, $x \preccurlyeq z$. If $x \sim z$ then $x <^1 y$, which contradicts not $x <^1 y$. If $x < z$ then $x < y$, which implies $x <^1 y$, a contradiction.
2. (x, y^*, z). Then $y < x$, $z \preccurlyeq y$, $x <^1 z$. From the last of these, $(x \sim t, t < z)$, which along with $(y < x, z \preccurlyeq y)$ contradicts $p10$.
3. (x^*, y, z). Then $y \preccurlyeq x$, not $y <^1 z$, $x < z$. Similar to Case 1.
4. (x, y^*, z^*). Then $y < x$, not $y <^2 z$, $x \preccurlyeq z$. If $x \sim z$ then $y <^2 z$, and if $x < z$ then $y < z$ and hence $y <^2 z$, contradicting not $y <^2 z$.
5. (x^*, y, z^*). Then $y \preccurlyeq x$, $z < y$, $x <^2 z$. From the last of these, $(x < t, t \sim z)$, which along with $(y \preccurlyeq x, z < y)$ contradicts $p10$.
6. (x^*, y^*, z). Then not $x <^2 y$, $z \preccurlyeq y$, $x < z$. Similar to Case 4.

Assume that $X/\approx$ is countable. Then $Y \cup Y^*$ is countable and by Theorem 2.2 there is a real-valued function f on $Y \cup Y^*$ such that, for all $b, c \in Y \cup Y^*$,

$$b <^3 c \Leftrightarrow f(b) < f(c).$$

For $x \in Y$ let $u(x) = f(x)$ and $\sigma(x) = f(x^*) - f(x)$. Then, using (2.17), $x < y \Leftrightarrow u(x) + \sigma(x) < u(y)$, for all $x, y \in Y$. Since $x <^3 x^*$ by (2.18), $\sigma > 0$. Let $u(x) = u(y)$ and $\sigma(x) = \sigma(y)$ whenever $x \approx y$ and $y \in Y$. Then (2.14) follows from Theorem 2.6 and Theorem 2.3(*d*). ◆

Semiorders

On adding *p*11 to *p*2 and *p*10 we obtain the following extension of Theorem 2.6.

THEOREM 2.8. *Suppose* $<$ *on* X *is a semiorder and, with* $<^1$ *and* $<^2$ *defined by* (2.12) *and* (2.13), $<^0$ *on* X *is defined by*

$$x <^0 y \Leftrightarrow x <^1 y \quad \text{or} \quad x <^2 y, \qquad \text{for all } x, y \in X. \tag{2.19}$$

Then $<^0$ *on* X *is a weak order.*

Proof. Asymmetry. $(x <^1 y, y <^1 x)$ and $(x <^2 y, y <^2 x)$ are prohibited by Theorem 2.6. Suppose $(x <^1 y, y <^2 x)$. Then $(x \sim z, z < y)$ and $(y < w, w \sim x)$ for some $z, w \in X$, which violates *p*11.

Negative Transitivity. By (2.19), not $x <^0 y \Rightarrow$ (not $x <^1 y$, not $x <^2 y$) and not $y <^0 z \Rightarrow$ (not $y <^1 z$, not $y <^2 z$). Therefore, by the negative transitivity of $<^1$ and $<^2$, (not $x <^1 z$, not $x <^2 z$), so that not $x <^0 z$ by (2.19). ◆

When $X/\approx$ is finite and $<$ is a semiorder, it is possible to make σ in (2.14) constant on X. A constructive proof of this is given in Scott and Suppes (1958) or in Suppes and Zinnes (1963). An alternative proof, similar to that given by Scott (1964), uses the Theorem of The Alternative which will be introduced in Chapter 4. Exercise 4.18 gives an outline of the alternative proof of the following theorem.

THEOREM 2.9. *Suppose* $<$ *on* X *is a semiorder and* $X/\approx$ *is finite. Then there is a real-valued function* u *on* X *such that*

$$x < y \Leftrightarrow u(x) + 1 < u(y), \qquad \text{for all } x, y \in X. \tag{2.20}$$

With an appropriate change in u, any positive number could be used in (2.20) in place of 1.

2.5 SUMMARY

A binary relation on a set is a weak order if it is asymmetric and negatively transitive. Defining indifference $\sim$ as the absence of strict preference, $\sim$ on X is an equivalence (reflexive, symmetric, transitive) when $<$ on X is a weak

order. If the set $X/\sim$ of equivalence classes of X under $\sim$ is countable when $<$ is a weak order then utilities $u(x), u(y), \ldots$ can be assigned to the elements in X so that $x < y \Leftrightarrow u(x) < u(y)$. This gives $x \sim y \Leftrightarrow u(x) = u(y)$ also.

The preference relation is a strict partial order when it is irreflexive and transitive. In this case indifference may be intransitive but $\approx$, defined by $x \approx y \Leftrightarrow (x \sim z \Leftrightarrow y \sim z$, for all $z \in X)$, is an equivalence. When $<$ on X is a strict partial order and $X/\approx$ is countable, utilities can be assigned so that $u(x) < u(y)$ if $x < y$, and $u(x) = u(y)$ if $x \approx y$.

Interval orders and semiorders lie between strict partial orders and weak orders. When $<$ on X is an interval order or a semiorder and $X/\approx$ is countable, we get $x < y \Leftrightarrow I(x)$ is wholly to the left of $I(y)$, where I is a function that assigns an interval of real numbers to each $x \in X$. If $<$ is a semiorder and $X/\approx$ is finite then all indifference intervals can be made to have the same length.

INDEX TO EXERCISES

1. Denumerable sets. 2. Binary relations. 3. Weak order. 4. Quasi order. 5–7. Asymmetric transitive closure. 8. Equivalence. 9. Partitions. 10–13. Interval orders and semiorders. 14. Choice sets. 15–16. Cartesian products. 17–18. Lexicographic orders. 19. Theorem 2.2. 20–21. Sets and relations.

Exercises

1. Prove that the following sets are denumerable: (*a*) $\{2, 4, 6, \ldots\}$, the set of all positive, even integers; (*b*) $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$; (*c*) the set of all positive rational numbers (Hint: place these in a two-dimensional array with 1/1, 1/2, 1/3, . . . in the first row, 2/1, 2/2, 2/3, . . . in the second row, and so forth); (*d*) the set of all rational numbers.

2. With Y the set of all living people, identify the meaning of cases (1) through (4) in Section 2.1 and state which of properties $p1$ through $p9$ hold for the binary relation identified as:

   a. "is a blood-line descendant of,"
   b. "is married to" (assuming monogamy throughout society),
   c. "is married to" (admitting polygamy),
   d. "is as old as,"
   e. "has fathered or mothered the same number of children as."

3. Suppose $\preccurlyeq$ on X is transitive and connected, and $\prec$ and $\sim$ are defined as follows: $x \prec y \Leftrightarrow$ not $y \preccurlyeq x$; $x \sim y \Leftrightarrow (x \preccurlyeq y, y \preccurlyeq x)$. Prove that $\prec$ is a weak order and that $\sim$ is an equivalence.

4. $\preccurlyeq$ on X is a *quasi order* iff it is reflexive and transitive. Prove that if $\preccurlyeq$ on X is a quasi order and $\prec$, $\sim$ are defined as in Exercise 3 then

a. $\sim$ on X is an equivalence

b. $\prec$ on X is a strict partial order

c. $(x \sim y, y \prec z) \Rightarrow x \prec z$, $(x \prec y, y \sim z) \Rightarrow x \prec z$.

5. If $\prec$ on X is a binary relation, the *transitive closure* $\prec^t$ of $\prec$ is defined as follows:

$$x \prec^t y \Leftrightarrow x \prec y \text{ or there are } x_1, x_2, \ldots, x_m \in X \text{ such that}$$
$$x \prec x_1, x_1 \prec x_2, \ldots, x_{m-1} \prec x_m, x_m \prec y.$$

Prove that if $\prec^t$ is asymmetric then $\prec^t$ is a strict partial order.

6. (*Continuation.*) Suppose X is countable. Use Theorem 2.5 to prove that there is a real-valued function u on X that satisfies (2.10) if and only if the transitive closure of $\prec$ is asymmetric.

7. (*Continuation.*) Give an example of a $\prec$ on X whose transitive closure is asymmetric and with u satisfying (2.10) and $\approx$ defined by (2.6) it is not possible for u to satisfy (2.11) also.

8. Using (2.2) and (2.6) prove that $\approx$ is an equivalence when $\prec$ on X is asymmetric.

9. A *partition* of a set Y is a set of nonempty subsets of Y such that each $x \in Y$ is in exactly one element of the partition. Prove that any partition of Y is a set of equivalence classes under some equivalence relation on Y.

10. Prove that $(p2, p10) \Rightarrow p6$ (transitivity) and that $(p2, p11) \Rightarrow p6$.

11. (From Fred Roberts.) $(X, \sim)$ is an *interval graph* $\Leftrightarrow$ a real interval $I(x)$ can be assigned to each $x \in X$ so that, for all $x, y \in X$, $x \sim y$ if and only if $I(x)$ and $I(y)$ intersect. Prove that if X is countable then $\prec$ on X is an interval order if and only if $\prec$ is transitive and $(X, \sim)$ is an interval graph.

12. (*Continuation.*) Roberts (1969). Suppose X is finite. Prove that $\prec$ on X is a semiorder if and only if $(X, \sim)$ is an interval graph and $(x \sim w, y \sim w, z \sim w$, not $x \sim y$, not $y \sim z$, not $x \sim z)$ is false whenever x, y, z, and w are in X.

13. Show that if (2.20) holds when $\prec$ is a semiorder and X is finite, then for any $\alpha > 0$ there is a real-valued function v_α on X such that, for all $x, y \in X$, $x \prec y \Leftrightarrow v_\alpha(x) + \alpha < v_\alpha(y)$.

14. Arrow (1959). Let F (the choice function) be a function that, for every nonempty subset Y of X, assigns a nonempty subset of Y to Y, so that $F(Y) \subseteq Y$ and $F(Y) \neq \varnothing$ for every $Y \subseteq X$ such that $Y \neq \varnothing$. Consider the following conditions on F.

TRANSITIVITY. $y \in F(\{x, y\}), z \in F(\{y, z\}) \Rightarrow z \in F(\{x, z\})$.

EXTENSION. $F(Y) = \{x : x \in Y \text{ and } x \in F(\{x, y\}) \text{ for every } y \in Y\}$, provided that the set $\{x : \cdots\}$ is not empty.

TE. If $x, y \in Y$, $x, y \in Y^*$, $x \in F(Y)$, and $y \notin F(Y)$ then $y \notin F(Y^*)$.

Interpret each of these conditions in your own words when $F(Y)$ is the individual's

set of most preferred elements in Y, that is his choice set. Then suppose that X is finite and prove that Transitivity and Extension hold if and only if TE holds. In doing this it may help to note that, with $x \preccurlyeq y \Leftrightarrow y \in F(\{x, y\})$, $\preccurlyeq$ is transitive and connected when Transitivity holds, and that, when the first two conditions hold then $F(Y) = \{x : x \in Y \text{ and } y \preccurlyeq x \text{ for all } y \in Y\}$.

15. Show that $\{(x_1, x_2) : x_1 \text{ and } x_2 \text{ are positive integers}\}$ is denumerable.

16. (*Continuation.*) The *Cartesian product* of sets X_1 and X_2 is $X_1 \times X_2 = \{(x_1, x_2) : x_1 \in X_1 \text{ and } x_2 \in X_2\}$. Use the preceding results to show that $X_1 \times X_2$ is denumerable if both X_1 and X_2 are denumerable.

17. (*Continuation.*) With $X = X_1 \times X_2$ let $X_1 = \{1, 2, \ldots\}$ and let X_2 be the set of all rational numbers between 0 and 1 inclusive. Define $\prec$ on X by $(x_1, x_2) \prec (y_1, y_2) \Leftrightarrow x_1 < y_1$ or $(x_1 = y_1, x_2 < y_2)$. (This weak order is a *lexicographic order* since it orders the pairs of numbers like two-letter words would be ordered in a dictionary.) Write out an explicit formula for u on $X_1 \times X_2$ that satisfies (2.5).

18. (*Continuation.*) Let $\prec$ be defined as in the preceding exercise, except that X_1 is the rationals between 0 and 1 inclusive and $X_2 = \{1, 2, \ldots\}$, the positive integers. Theorem 2.2 says that there is a real-valued u on $X = X_1 \times X_2$ that satisfies (2.5). Can you write out an explicit formula for u on $X_1 \times X_2$ that satisfies (2.5)? If not, explain why not.

19. Prove Theorem 2.2 when $X/\sim$ is finite.

20. Let $A \subseteq B$ mean that A is a subset of B and $A = B$ if and only if $A \subseteq B$ and $B \subseteq A$. $A \cup B$ is the set of all elements in A or in B, and $A \cap B$ is the set of all elements in both A and B. Let $\varnothing$ denote the empty set (set with no elements). With Y a set, let $\Delta = \{(x, x) : x \in Y\}$; if R is a binary relation on Y let $R' = \{(y, x) : (x, y) \in R\}$; if R and S are binary relations on Y, let $RS = \{(x, z) : xRy \text{ and } ySz \text{ for some } y \in Y\}$. Express $p1$ through $p11$ of the chapter in terms of these definitions. For example, $p1$ can be written as $\Delta \subseteq R$.

21. (*Continuation.*) Verify that when the given sets are binary relations on a set Y, then

a. $\Delta' = \Delta$; $\varnothing' = \varnothing$ [$\varnothing$ is the empty binary relation]
b. $(A \cup B)' = A' \cup B'$; $(A \cap B)' = A' \cap B'$
c. $(AB)C = A(BC)$
d. $A\varnothing = \varnothing A = \varnothing$
e. $A\Delta = \Delta A = A$
f. $A \subseteq B$ and $C \subseteq D$ imply $AC \subseteq BD$.

(See Chipman (1960) for additional material of this kind.)

1

Chapter 3

UTILITY THEORY FOR UNCOUNTABLE SETS

This chapter extends the theory of preference-preserving utility functions to include uncountable sets. A new condition of order denseness is used for this purpose. After proving basic theorems for weak orders and strict partial orders we shall consider preferences on subsets of n-dimensional Euclidean space. The chapter concludes with a discussion of continuous utility functions.

An *uncountable* set is a set that is not countable: it is neither finite nor denumerable. The following examples introduce some other new terms.

1. The set of all real numbers is uncountable. This set, denoted by Re or E^1, is *one-dimensional Euclidean space.* The intervals of numbers $[a, b] = \{x : a \leq x \leq b\}$, $[a, b) = \{x : a \leq x < b\}$, $(a, b] = \{x : a < x \leq b\}$, and $(a, b) = \{x : a < x < b\}$ are uncountable when $a < b$. $[a, b]$ is a *closed interval:* (a, b) is an *open interval.* (a, b) is also used to denote an ordered pair of elements. The context should clarify the usage.

2. The set $\{(x_1, x_2, \ldots, x_n) : x_i \in \text{Re for } i = 1, \ldots, n\}$, denoted as Re^n or E^n and called *n-dimensional Euclidean space*, is uncountable. E^2 is the real plane. In the vector $(x_1, x_2, \ldots, x_n)$, the ith *component* is x_i.

3. The set $\{(x_1, x_2, \ldots) : x_i \in \{0, 1\} \text{ for } i = 1, 2, \ldots\}$ is uncountable. Although $\{(x_1, x_2, \ldots, x_n) : x_i \in \{0, 1\} \text{ for } i = 1, 2, \ldots, n\}$ is finite for each n, the given denumerable-dimensional set is uncountable (and not denumerable). On the other hand, $\{(x_1, x_2) : x_i \in \{1, 2, \ldots\} \text{ for } i = 1, 2\}$ is denumerable.

3.1 THE DENSENESS AXIOM AND WEAK ORDERS

We shall now extend Theorem 2.2 to cover the case where $X/\sim$ may not be countable. To do this we shall introduce an assumption concerning the concept of order denseness.

Definition 3.1. Let R be a binary relation on a set Y. Then $Z \subseteq Y$ is *R-order dense* in Y if and only if, whenever xRy and x and y are in Y but not Z, there is a $z \in Z$ such that (xRz, zRy).

Since there is a rational number between any two distinct real numbers, the countable set of rational numbers is $<$-order dense in Re. For the following theorem, $<'$ on $X/\sim$ is defined by (2.4).

THEOREM 3.1. *There is a real-valued function u on X such that*

$$x < y \Leftrightarrow u(x) < u(y), \qquad \textit{for all } x, y \in X, \tag{3.1}$$

if and only if $<$ on X is a weak order and there is a countable subset of $X/\sim$ that is $<'$-order dense in $X/\sim$.

Unfortunately, the countable order denseness condition does not have a simple, intuitive interpretation. To see how this condition can fail, suppose $X = \mathrm{Re}^2$ with $<$ the lexicographic order

$$(x_1, x_2) < (y_1, y_2) \Leftrightarrow x_1 < y_1 \quad \text{or} \quad (x_1 = y_1, x_2 < y_2).$$

Then $X/\sim = \{\{x\}: x \in X\}$, so that $\{x\} <' \{y\} \Leftrightarrow x < y$. With x_1 fixed it takes a denumerable subset of Re to obtain an $<$-order dense subset on $\{x_1\} \times \mathrm{Re}$. But there is an uncountable number of such x_1 and it follows that no countable subset of Re^2 is $<$-order dense in Re^2.

For another example let $X = [-1, 1]$. The *absolute value* of x, written $|x|$, is defined by $|x| = x$ if $x \geq 0$, $|x| = -x$ if $x < 0$. Define $<$ on X by

$$x < y \Leftrightarrow |x| < |y| \quad \text{or} \quad (|x| = |y|, x < y).$$

Suppose Y is $<$-order dense in $[-1, 1]$. With $x \in (0, 1]$, $-x < x$ and there is no y with $|y| \neq |x|$ such that $-x < y < x$. Hence either $-x$ or x must be in Y for each $x \in (0, 1]$. Thus, every $<$-order dense subset Y of $[-1, 1]$ contains a subset that is in one-to-one correspondence with $(0, 1]$, which is uncountable.

Proof of Theorem 3.1

Before proving the theorem, several additional notions will be defined. If A and B are sets, the union $A \cup B$ of A and B is the set of all elements in A or B. The *relative difference* $A - B$ is the set of all elements in A but not B.

Let A be a set of numbers all of which are less than some number not in A. Then the *least upper bound* or *supremum* of A is the smallest number that is as large as every number in A:

$$\sup A = \text{smallest } y \text{ such that } x \leq y \text{ for all } x \in A.$$

If all numbers in A exceed some number not in A then the *greatest lower*

bound or *infimum* of A is the largest number that is as small as every number in A:

$$\inf A = \text{largest } y \text{ such that } y \leq x \text{ for all } x \in A.$$

For example, $\sup\{1, 2, 3\} = 3$, $\inf\{1, 2, 3\} = 1$, $\sup(0, 1) = 1$ and $\inf(0, 1) = 0$. In the last two cases sup and inf are not in A.

Proof of Necessity. Let (3.1) hold. Then $<$ on X must be a weak order, and $<'$ on $X/\sim$ is a strict order, with $a <' b \Leftrightarrow u(a) < u(b)$, where $u(a) = u(x)$ whenever $x \in a$. Let $\mathcal{C}$ be the denumerable set of closed intervals in Re with distinct, rational endpoints. For each $I \in \mathcal{C}$ that contains some $u(a)$ for $a \in X/\sim$, select one such a. Let A be the subset of $X/\sim$ thus selected. A is countable. Next, let

$$K = \{(b, c): b, c \in X/\sim - A, b <' c, b <' a <' c \text{ for no } a \in A\}.$$

If $(b, c) \in K$, then $b <' a <' c$ for no $a \in X/\sim$, for otherwise there would be a $d \in A$ with $b <' d <' c$ since for every point in the open interval $(u(b), u(c))$ there is an $I \in \mathcal{C}$ that includes the point with $I \subset (u(b), u(c))$. Hence no two open intervals $(u(b), u(c))$ for $(b, c) \in K$ overlap, so that K must be countable. Therefore,

$$B = \{b: b \in X/\sim, \text{ there is a } c \in X/\sim \text{ such that } (b, c) \in K \text{ or } (c, b) \in K\}$$

is countable and hence $A \cup B$ is countable. Moreover, if $b, c \in X/\sim - A \cup B$ and $b <' c$, then there is an $a \in A \cup B$ such that $b <' a <' c$. Thus the countable order denseness condition is necessary for (3.1). ◆

Proof of Sufficiency. We assume that $<$ on X is a weak order and will work with the strict order $<'$ on $X/\sim$. We shall assume that A includes the least and/or most preferred ($<'$) elements in $X/\sim$, if such exist, and that A is countable and is $<'$-order dense in $X/\sim$. Let

$$B = \{b: b \in X/\sim - A, \text{ either } \{a: a \in A, b <' a\} \text{ has a least preferred element } a_b \text{ or } \{c: c \in A, c <' b\} \text{ has a most preferred element } c_b\}.$$

With $b \in X/\sim - A$, $\{a: a \in A, b <' a\}$ and $\{c: c \in A, c <' b\}$ are two disjoint subsets of A whose union equals A. It follows that a given $a \in A$ can be an a_b for at most one $b \in X/\sim - A$, and that a given $c \in A$ can be a c_b for at most one $b \in X/\sim - A$. Hence B is countable and therefore

$$C = A \cup B$$

is countable. Moreover,

1. *There is no least preferred* $a \in \{a: a \in C, b <' a\}$ *for any* $b \in X/\sim - C$
2. *There is no most preferred* $c \in \{c: c \in C, c <' b\}$ *for any* $b \in X/\sim - C$.

For proof, suppose (1) is false and a_b is the least preferred element in $\{a:a \in C, b <' a\}$ for some $b \in X/\sim - C$. Then a_b cannot be in A, for otherwise $b \in B$. But then $c <' b <' a_b <' a$ for all $c \in \{c:c \in A, c <' b\}$ and all $a \in \{a:a \in A, b <' a\}$ and there is no element in A between b and a_b, in violation of the order denseness assumption. Hence (1) is true and, by a symmetric proof, (2) is true.

By the proof of Theorem 2.2 there is a real-valued function u on C such that $a <' c \Leftrightarrow u(a) < u(c)$, for all $a, c \in C$. For each $b \in X/\sim - C$ let

$$u^b = \{u(a):a \in C, b < a\}$$

$$u_b = \{u(c):c \in C, c < b\}$$

and set

$$u(b) = \tfrac{1}{2}(\sup u_b + \inf u^b), \tag{3.2}$$

where, since $u(c) < u(a)$ for all $c \in u_b$ and $a \in u^b$, $\sup u_b \leq \inf u^b$. From (2) and (1) above it follows that for each $b \in X/\sim - C$,

$$u(c) < \sup u_b, \qquad \text{for all } u(c) \in u_b$$

$$\inf u^b < u(a), \qquad \text{for all } u(a) \in u^b.$$

Hence $u(c) < u(b) < u(a)$ for all $c \in \{c:c \in C, c <' b\}$ and all $a \in \{a:a \in C, b <' a\}$. Hence $u(b) \neq u(a)$ when $b \in X/\sim - C$ and $a \in C$, and the extension of u by (3.2) preserves the ordering of the $b \in X/\sim - C$ and the $a \in C$.

Suppose then that $b, c \in X/\sim - C$. If $b <' c$ then $b <' a <' c$ for some $a \in C$ so that $u(b) < u(a)$ and $u(a) < u(c)$ and hence $u(b) < u(c)$. Conversely, if $u(b) < u(c)$, there is, by definition of supremum and (1), a $u(a) \in u^b$ such that $u(b) < u(a) < u(c)$, which yields $b <' a$ and $a <' c$ and therefore $b <' c$ by transitivity. Hence, for all $a, b \in X/\sim$, $a <' b \Leftrightarrow u(a) < u(b)$. Defining $u(x) = u(a)$ when $x \in a$, (3.1) follows. ◆

The above proof is patterned after outlines in Birkhoff (1948, p. 32) and Luce and Suppes (1965, pp. 263–264). Our proof is similar also to Debreu's proof of his Lemma II (1954, pp. 161–162).

3.2 PREFERENCE AS A STRICT PARTIAL ORDER

We shall now consider an appropriate generalization of Theorem 2.5 for strict partial orders. Throughout this section $<^*$ on $X/\approx$ is defined as in (2.7) with $\approx$ as in (2.6).

THEOREM 3.2. *Suppose $<$ on X is a strict partial order and there is a countable subset of $X/\approx$ that is $<^*$-order dense in $X/\approx$. Then there is a*

real-valued function u on X such that

$$x \prec y \Rightarrow u(x) < u(y), \qquad \textit{for all } x, y \in X, \tag{3.3}$$

$$x \approx y \Rightarrow u(x) = u(y), \qquad \textit{for all } x, y \in X. \tag{3.4}$$

In this case the denseness condition is not necessary for (3.3) and (3.4). Suppose for example that $X = \text{Re}$ and define

$$x \prec y \Leftrightarrow x < y \quad \text{and} \quad y = x + n \qquad \text{for some positive integer } n.$$

Then $u(x) = x$ satisfies (3.3) and (3.4), and $X/\approx\, = \{\{x\} : x \in X\}$. If $Z \subseteq X$ is countable then there is an x such that neither x nor $x + 1$ is in Z. But with $x \prec x + 1$, there is no $z \in Z$ such that $x \prec z \prec x + 1$. It follows that there is no countable subset of $X/\approx$ that is $\prec^*$-order dense in $X/\approx$.

Our proof of Theorem 3.2 is based on an ingenious proof of a somewhat more general theorem given by Richter (1966).

Proof of Theorem 3.2. Let the hypotheses of the theorem hold. By Theorem 2.3, $\prec^*$ on $X/\approx$ is a strict partial order. Let A be a countable subset of $X/\approx$ that is $\prec^*$-order dense in $X/\approx$. By Theorem 2.4 there is a strict order $\prec^0$ on $X/\approx$ that includes $\prec^* : a \prec^* b \Rightarrow a \prec^0 b$. Define a binary relation E on $X/\approx$ as follows:

$$aEb \Leftrightarrow a = b \quad \text{or} \quad (a, b \notin A \text{ and } a \prec^0 c \prec^0 b \text{ or } b \prec^0 c \prec^0 a \text{ for no } c \in A).$$

Then E is obviously reflexive and symmetric and is in fact an equivalence on $X/\approx$. For transitivity suppose (aEb, bEc) with $a \neq b \neq c \neq a$ (to avoid the trivial cases). If $(a \prec^0 b, b \prec^0 c)$ or $(a \prec^0 b, c \prec^0 b)$ or $(b \prec^0 a, b \prec^0 c)$ or $(b \prec^0 a, c \prec^0 b)$, which are the only four possibilities, then there is no $d \in A$ such that $a \prec^0 d \prec^0 c$ or $c \prec^0 d \prec^0 a$. Hence aEc.

Let r, s, and t be equivalence classes in the set of such classes in $X/\approx$ under E. That is, $r \in (X/\approx)/E$. Define $\prec^1$ on these classes as follows:

$$r \prec^1 s \Leftrightarrow r \neq s \quad \text{and} \quad a \prec^0 b \qquad \text{for some (and thus for all) } a \in r, b \in s.$$

Since $\prec^0$ on $X/\approx$ is a strict order and E on $X/\approx$ is an equivalence, $\prec^1$ on $X/\approx/E$ is a strict order. Moreover, $B = \{r : r \in X/\approx/E \text{ and } a \in r \text{ for some } a \in A\}$ is $\prec^1$-order dense in $X/\approx/E$. For suppose r, s are not in B and $r \prec^1 s$. Then, with $a \in r$ and $b \in s$, $a \prec^0 b$ and $a, b \notin A$. Since not aEb there must be a $c \in A$ such that $a \prec^0 c \prec^0 b$. With $c \in t$ it follows that $t \in B$ and $r \prec^1 t \prec^1 s$.

It then follows from the proof of Theorem 3.1 that there is a real-valued function f on $X/\approx/E$ such that

$$r \prec^1 s \Leftrightarrow f(r) < f(s), \qquad \text{for all } r, s \in X/\approx/E. \tag{3.5}$$

Suppose that, with $a \in r$ and $b \in s$, $a \prec^* b$. Then $a \prec^0 b$. Therefore either

$r = s$ or $r <^1 s$. If either a or b is in A then $r \neq s$ since $a \neq b$ and hence not aEb. If $a, b \notin A$ and $r = s$ then $a <^0 c <^0 b$ for no $c \in A$, which is false since A is $<^*$-order dense in $X/\approx$ and hence if $a <^* b$ and $a, b \notin A$ then $a <^* c <^* b$ (and thus $a <^0 c <^0 b$) for some $c \in A$. Therefore $a <^* b \Rightarrow r <^1 s$. Defining $u(a) = f(r)$ when $a \in r$ it follows from (3.5) that if $a <^* b$ then $u(a) < u(b)$. Defining $u(x) = u(a)$ when $x \in a$ and observing that if $x < y$ and $(x \in a, y \in b)$ then $a <^* b$, it follows that $x < y \Rightarrow u(x) < u(y)$. It is clear also that $u(x) = u(y)$ when $x, y \in a$. ◆

3.3 PREFERENCES ON Re^n

Preferences in many decision situations are influenced by multiple factors. Hence a large part of our study will focus on sets whose elements are n-tuples. When the components of the n-tuples are real numbers, the n-tuples are called vectors.

This section looks at the special case where X equals Re^n or is a rectangular subset of Re^n, by which is meant the Cartesian product of n real intervals, including perhaps infinite intervals such as (a, ∞), the set of all numbers greater than a, and $(-\infty, \infty) = \text{Re}$.

When $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$ are vectors in Re^n and α, β are scalars (real numbers), we define *multiplication by scalars* and *vector addition* by

$$\begin{aligned} \alpha x + \beta y &= (\alpha x_1, \ldots, \alpha x_n) + (\beta y_1, \ldots, \beta y_n) \\ &= (\alpha x_1 + \beta y_1, \ldots, \alpha x_n + \beta y_n). \end{aligned} \quad (3.6)$$

After illustrating a utility function for increasing preferences in two dimensions we shall consider some formal theory for such cases.

Example

We consider preferences of the president of a company on a set of two-dimensional vectors (x_1, x_2) where x_1 denotes net profit for the coming year and x_2 denotes the company's market share for the coming year. $X_1 = [-\$5 \text{ million}, \$5 \text{ million}]$ and $X_2 = [10\%, 30\%]$. A utility surface that might reflect the president's preferences is shown in Figure 3.1. If $<$ on $X_1 \times X_2$ is a weak order and (3.1) holds then all $(x_1, x_2) \in X_1 \times X_2$ with equal utility constitute an element in $X/\sim$. These equivalence classes are variously called *indifference curves*, *trade-off curves*, *indifference loci*, *isoutility contours*, and so forth. The family of indifference curves in the plane constitutes an *indifference map*. Two curves of the indifference map are illustrated in the figure.

If indifference were not transitive in this example then the preceding interpretation for an element in $X/\sim$ does not apply.

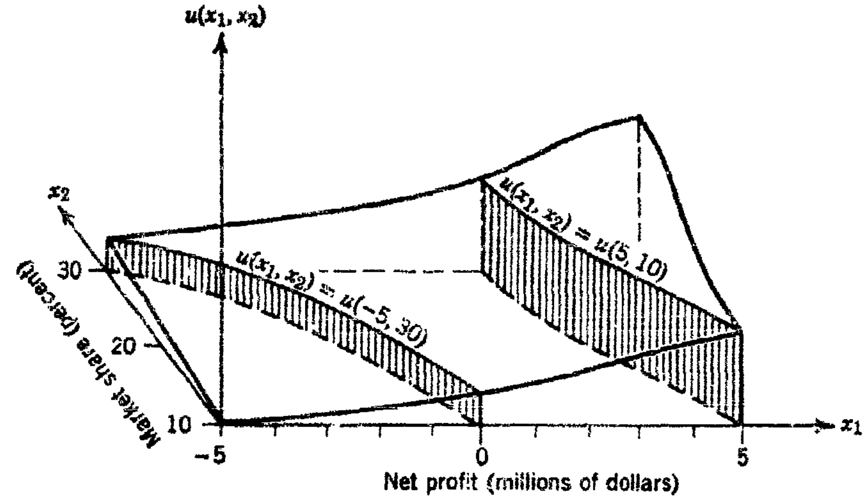

Figure 3.1 Unidimensional utilities on a two-dimensional space.

Increasing Preferences with Weak Orders

Let X_i, $i = 1, 2, \ldots, n$ be nonempty sets. Their Cartesian product is $X_1 \times X_2 \times \cdots \times X_n = \{(x_1, x_2, \ldots, x_n): x_i \in X_i \text{ for } i = 1, 2, \ldots, n\}$. In this subsection we assume that each X_i is an interval of real numbers, so that $X = X_1 \times \cdots \times X_n$ is a rectangular subset of Re^n. Elements in X_i could be amounts of money allocated to activity i or earned in year i, or they could be amounts of commodity i purchased during a fixed time period, and so forth.

With $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$, we *define* $x < y \Leftrightarrow x \neq y$ *and* $x_i \leq y_i$ *for* $i = 1, \ldots, n$.

THEOREM 3.3. *Suppose that X is a rectangular subset of* Re^n *and that the following hold throughout X:*

1. $\prec$ *on X is a weak order,*
2. $x < y \Rightarrow x \prec y$,
3. $(x \prec y, y \prec z) \Rightarrow \alpha x + (1 - \alpha)z \prec y$ *and* $y \prec \beta x + (1 - \beta)z$ *for some* $\alpha, \beta \in (0, 1)$.

Then there is a real-valued function u on X that satisfies (3.1).

The second condition (monotonicity, nonsaturation, nonsatiety, dominance, etc.) states that preference increases with any increase in quantity. Condition 3 is an Archimedean condition that will be used to establish a countable order dense subset. For the third condition to hold it may be necessary in some cases to have α very near to 1 and β very near to zero.

In proving the theorem we shall first prove the following lemma.

LEMMA 3.1. *The hypotheses of Theorem 3.3 imply that if $x, y, z \in X$ and $x < y < z$, then $y \sim \alpha x + (1 - \alpha)z$ for exactly one $\alpha \in (0, 1)$.*

Proof. If $y \sim \alpha x + (1 - \alpha)z$ for no $\alpha \in (0, 1)$, it follows from the hypotheses that there is a $\beta \in (0, 1)$ such that either

$$y \prec \alpha x + (1 - \alpha)z \qquad \text{for all } \alpha \leq \beta \tag{3.7}$$

$$\alpha x + (1 - \alpha)z \prec y \qquad \text{for all } \alpha > \beta \tag{3.8}$$

or

$$y \prec \alpha x + (1 - \alpha)z \qquad \text{for all } \alpha < \beta \tag{3.9}$$

$$\alpha x + (1 - \alpha)z \prec y \qquad \text{for all } \alpha \geq \beta. \tag{3.10}$$

We consider the latter case. By (3.10) and the hypotheses, $\beta x + (1 - \beta)z \prec y \prec z$. Hence, by condition 3 of Theorem 3.3, there is an $\alpha \in (0, 1)$ such that $\alpha[\beta x + (1 - \beta)z] + (1 - \alpha)z \prec y$, or $\alpha\beta x + (1 - \alpha\beta)z \prec y$. But since $\alpha\beta < \beta$, (3.9) says that $y \prec \alpha\beta x + (1 - \alpha\beta)z$, a contradiction. Hence (3.9) and (3.10) can't hold. A similar proof shows that (3.7) and (3.8) can't hold. Hence $y \sim \alpha x + (1 - \alpha)z$ for some $\alpha \in (0, 1)$. If $y \sim \alpha_1 x + (1 - \alpha_1)z$ and $y \sim \alpha_2 x + (1 - \alpha_2)z$ then $\alpha_1 x + (1 - \alpha_1)z \sim \alpha_2 x + (1 - \alpha_2)z$ by the transitivity of $\sim$, which can only be true if $\alpha_1 = \alpha_2$: for if $\alpha_1 < \alpha_2$ then $\alpha_2 x + (1 - \alpha_2)z < \alpha_1 x + (1 - \alpha_1)z$ since $x < z$. ◆

Proof of Theorem 3.3. In view of Theorem 3.1 we need to show that $X/\sim$ contains a countable subset that is $\prec'$-order dense in $X/\sim$.

Let Y_i be the set of all rational numbers plus any finite end point at any closed end of X_i (if such exist). Let $Z_i = X_i \cap Y_i$. Z_i is countable. Let $W_i = \{\alpha x_i + (1 - \alpha)y_i : \alpha \text{ is a rational number in } [0, 1] \text{ and } x_i, y_i \in Z_i\}$. W_i is a countable set. Let $W = W_1 \times W_2 \times \cdots \times W_n$. W is countable. Let A consist of all elements in $X/\sim$ that contain one or more elements in W. A is countable since any $x \in W$ is in exactly one $a \in X/\sim$. Suppose $a, b \in X/\sim - A$ with $a \prec' b$. We need to show that there is a $c \in A$ such that $a \prec' c \prec' b$. To do this it will suffice to show that when $x, y \in X - W$ and $x \prec y$ then there is a $z \in W$ such that $x \prec z \prec y$. We consider two cases as follows.

Case 1: $x < y$. Then there are $z^1, z^2 \in Z_1 \times \cdots \times Z_n$ such that $z^1 < x$ and $y < z^2$. Lemma 3.1, weak order, and condition 2 of the theorem imply that there are α, β with $0 < \alpha < \beta < 1$ such that $x \sim \beta z^1 + (1 - \beta)z^2$, $y \sim \alpha z^1 + (1 - \alpha)z^2$. Let γ be any rational number in the interval (α, β). Then, by weak order and condition 2, $x \prec \gamma z^1 + (1 - \gamma)z^2 \prec y$. Since $z^1, z^2 \in Z_1 \times \cdots \times Z_n$ and γ is rational, $\gamma z^1 + (1 - \gamma)z^2 \in W$.

Case 2: $x < y$ is false (with $x, y \in X - W$ and $x \prec y$). Let $v_i = \inf\{x_i, y_i\}$ and $w_i = \sup\{x_i, y_i\}$. Then $v < x < w$ and $v < y < w$. It follows that there are α, β with $0 < \alpha < \beta < 1$ such that $x \sim \beta v + (1 - \beta)w$ and

$y \sim \alpha v + (1 - \alpha)w$. Since $\beta v + (1 - \beta)w < \alpha v + (1 - \alpha)w$, it follows from the Case 1 proof that there is a $z \in W$ such that $\beta v + (1 - \beta)w < z < \alpha v + (1 - \alpha)w$. Hence $x \prec z \prec y$. ◆

If preference decreases rather than increases as $x_i \in X_i$ increases, Theorem 3.3 can still be used after a change of variable from x_i to $y_i = -x_i$.

Nondecreasing Preferences with Strict Partial Order

We conclude this section with a theorem that uses generally weaker conditions than those of the preceding theorem. We shall use the non-negative orthant $\{(x_1, \ldots, x_n): x_i \geq 0 \text{ for } i = 1, \ldots, n\}$ of Re^n. This is often used by mathematical economists in investigations of consumer preference or consumer choice. In this context the vectors are called commodity bundles.

$$x \ll y \quad \text{means that} \quad x_i < y_i \quad \text{for} \quad i = 1, \ldots, n.$$

THEOREM 3.4. *Suppose that X is the non-negative orthant of* Re^n *and that the following hold throughout X;*

1. $\prec$ *on X is a strict partial order,*
2. $[(x \ll y, y \prec z)$ *or* $(x \prec y, y \ll z)] \Rightarrow x \prec z$,
3. $x \prec y \Rightarrow z \prec y$ *for some z such that* $x \ll z$.

Then there is a real-valued function u on X that satisfies (3.3).

The notion of nondecreasing preferences comes from condition 2. Irreflexivity and condition 2 say that $x \ll y \Rightarrow$ not $y \prec x$: an increase in every commodity does not decrease preference. Condition 3 says that if y is preferred to x then increases (perhaps very slight) can be made in all components of x, and y will still be preferred to the augmented x.

Proof of Theorem 3.4. Let the hypotheses hold. Define $x \prec^1 y \Leftrightarrow x \prec y$ or $x \ll y$. Conditions 1 and 2 imply that $\prec^1$ is a strict partial order. From $\prec^1$ we can define $\sim^1$ and $\approx^1$ in the manner of (2.2) and (2.6). By Theorem 2.3, $\approx^1$ on X is an equivalence and $\prec^{1*}$ on $X/\approx^1$, defined in the manner of (2.7), is a strict partial order. To show that there is a countable subset of $X/\approx^1$ that is $\prec^{1*}$-order dense in $X/\approx^1$, it suffices to show that the set of rational vectors in X (all components rational) is $\prec^1$-order dense in X. Suppose then that x and y are not rational and $x \prec^1 y$. If $x \ll y$ then $x \ll z \ll y$ for some rational z, and hence $x \prec^1 z \prec^1 y$. If $x \prec y$ then, by condition 3, $z \prec y$ for some z such that $x \ll z$. Then $x \ll t \ll z$ for some rational t. By condition 2, $t \prec y$. Hence $x \prec^1 t \prec^1 y$. Therefore, by Theorem 3.2, there is a real-valued function u on X such that $x \prec^1 y \Rightarrow u(x) < u(y)$. Then $x \prec y \Rightarrow u(x) < u(y)$ since $x \prec y \Rightarrow x \prec^1 y$. ◆

3.4 CONTINUOUS UTILITIES

Continuity formalizes the intuitive notion that if two elements in X are not very different then their utilities should be close together. The difference between x and y can be thought of either in terms of their relative proximity under $\prec$ or in terms of a structure for X that is related to $\prec$ in some way.

Part of the interest in continuity stems from the fact that, when continuity holds, the utility function will attain a maximum value on a suitably restricted subset of X. Suppose for example that X is the non-negative orthant of Re^n and that an individual can spend his income $m \geq 0$ on the n commodities whose unit prices are $p_1 > 0, p_2 > 0, \ldots, p_n > 0$. His choice is restricted to $(p, m) = \{x : x \in X \text{ and } \sum_{i=1}^{n} p_i x_i \leq m\}$. If $\prec$ satisfies the conditions of Theorem 3.3 then there is a u that satisfies (3.1) and is continuous, and there is an $x^* \in (p, m)$ that satisfies $\sum p_i x_i^* = m$ and $\sup \{u(x) : x \in (p, m)\} = u(x^*)$. Or suppose that $\prec$ satisfies the conditions of Theorem 3.4. Then there is a u that satisfies (3.3) and is upper semicontinuous and there is an $x^* \in (p, m)$ such that $\sup \{u(x) : x \in (p, m)\} = u(x^*)$. (See, for example, Thielman (1953, p. 102).)

Definitions for Continuity

To consider a general definition of continuity we require the following notions. The *union* ($\cup$) of a set of subsets of X is the set of elements that appear in at least one of the subsets. The *intersection* ($\cap$) of a set of subsets of X is the set of elements that appear in every one of the subsets.

Definition 3.2. A *topology* $\mathcal{T}$ for a set X is a set of subsets of X such that

1. The empty set $\emptyset$ (which is always a subset of X) is in $\mathcal{T}$,
2. $X \in \mathcal{T}$,
3. The union of arbitrarily many sets in $\mathcal{T}$ is in $\mathcal{T}$,
4. The intersection of any finite number of sets in $\mathcal{T}$ is in $\mathcal{T}$.

If $\mathcal{T}$ is a topology for X, the pair $(X, \mathcal{T})$ is a *topological space. By definition,* the subsets of X in $\mathcal{T}$ are called *open sets.*

The usual topology $\mathcal{U}$ for Re is the set of open intervals along with their arbitrary unions and finite intersections. The *relative usual topology* for $X \subseteq \text{Re}$ is $\{A \cap X : A \in \mathcal{U}\}$. When $X = [0, 2]$, the closed interval $[0, 2]$ is an open set in the relative usual topology, but it is the only nonempty closed interval in X that is an open set in the relative usual topology.

Definition 3.3. If $(X, \mathcal{T})$ is a topological space then a real-valued function u on X is *continuous in the topology* $\mathcal{T}$ if and only if $A \in \mathcal{U} \Rightarrow \{x : x \in X, u(x) \in A\} \in \mathcal{T}$.

Suppose $X = [0, 2]$ and $\mathcal{T} = \{A \cap [0, 2]: A \in \mathcal{U}\}$. Then the function $u(x) = x$ for all $x \in X$ is continuous in $\mathcal{T}$, but the two-part function $f(x) = x$ for $x \in [0, 1]$ and $f(x) = x + 1$ for $x \in (1, 2]$ is not continuous because of its *gap* or jump at $x = 1$. For example, $(1/2, 3/2) \in \mathcal{U}$ but $\{x: x \in [0, 2], f(x) \in (1/2, 3/2)\} = (1/2, 1]$ is not in $\mathcal{T}$.

Necessary and Sufficient Conditions for Continuity

Assume that u on X satisfies (3.1) and is continuous in the topology $\mathcal{T}$. For any $y \in X$ the sets $\{b: b < u(y)\}$ and $\{a: u(y) < a\}$ are open sets in $\mathcal{U}$: hence $\{x: x \in X, x \prec y\}$ and $\{x: x \in X, y \prec x\}$ must be open sets in $\mathcal{T}$ for every $y \in X$. Again, if u is continuous in the topology $\mathcal{T}$ and if $x \prec y$, so that $u(x) < u(y)$, then there are open sets $A_x, A_y \in \mathcal{U}$ such that $u(x) \in A_x$ and $a < u(y)$ for every $a \in A_x$, and $u(y) \in A_y$ and $u(x) < b$ for every $b \in A_y$: hence there is an open set $\{z: u(z) \in A_x\}$ containing x such that $z \prec y$ for every z in this set and there is an open set $\{w: u(w) \in A_y\}$ containing y such that $x \prec w$ for every w in this set.

The foregoing paragraph sets forth two necessary conditions for continuity. Each condition is also sufficient for continuity.

THEOREM 3.5. *If* $(X, \mathcal{T})$ *is a topological space and there is a real-valued function on X satisfying* (3.1), *then there is a real-valued function on X satisfying* (3.1) *and continuous in the topology* $\mathcal{T}$ *if and only if*

1. $\{x: x \in X, x \prec y\} \in \mathcal{T}$ *and* $\{x: x \in X, y \prec x\} \in \mathcal{T}$ *for every* $y \in X$, or
2. *If* $x, y \in X$ *and* $x \prec y$, *then there are sets* $T_x, T_y \in \mathcal{T}$ *such that* $x \in T_x$, $y \in T_y$, $x' \prec y$ *for every* $x' \in T_x$ *and* $x \prec y'$ *for every* $y' \in T_y$.

Proof. The sufficiency of conditions 1 and 2 for continuity can be established by showing that 2 implies 1 and that 1 implies that some u satisfying (3.1) is continuous in $\mathcal{T}$.

Let y be any element in X. We show that condition 2 implies that $\{x: x \in X, x \prec y\} \in \mathcal{T}$; a symmetric proof suffices for the other part of condition 1. If $x \prec y$ for no $x \in X$ then $\{x: x \in X, x \prec y\} = \varnothing$, which is in $\mathcal{T}$. If $x \prec y$, then by condition 2 there is a set $T_x \in \mathcal{T}$ containing x such that $x' \prec y$ for all $x' \in T_x$. The union of all such T_x is $\{x: x \in X, x \prec y\}$, which is in $\mathcal{T}$ by part 3 of Definition 3.2.

To show that condition 1 implies that some u satisfying (3.1) is continuous in $\mathcal{T}$, we follow Debreu (1964). Let u on X satisfy (3.1), with $u(X) = \{u(x): x \in X\}$. A *gap* of $u(X)$ is a nonempty interval I in Re such that no point in $u(X)$ is in I and, with $a \in I$, $I = \{b: u(x) < b < u(y)$ for all $u(x) \in \{u(x): x \in X, u(x) < a\}$ and all $u(y) \in \{u(y): y \in X, a < u(y)\}\}$. Debreu's basic theorem (p. 285) asserts that, with u on X satisfying (3.1), there is a function

v on X that satisfies (3.1) such that all gaps of $v(X)$ are open intervals in $\mathcal{U}$. Debreu's proof of this (pp. 285–289) will not be repeated here.

Let v on X satisfy (3.1) with all gaps of $v(X)$ open. With $a \in \text{Re}$, let $(-\infty, a) \in \mathcal{U}$ be the open interval of all numbers less than a. If $a \in v(X)$ with $a = v(y)$, then $\{x : v(x) \in (-\infty, a)\} = \{x : x \prec y\}$ which by condition 1 is in $\mathcal{T}$. If $a \notin v(X)$ and a is in a gap of $v(X)$, this gap has the form (a_1, a_2) with $a \in (a_1, a_2)$ and $a_1, a_2 \in v(X)$: then, $\{x : v(x) \in (-\infty, a)\} = \{x : x \prec z\}$ where $a_2 = v(z)$, and again by condition 1 this set is in $\mathcal{T}$. Finally, if $a \notin v(X)$ and it is in no gap of $v(X)$, then either

1. $a \leq \inf v(X)$ so that $\{x : v(x) \in (-\infty, a)\} = \varnothing$, in $\mathcal{T}$, or
2. $\sup v(X) \leq a$ so that $\{x : v(x) \in (-\infty, a)\} = X$, in $\mathcal{T}$, or
3. $a = \sup \{v(x) : x \in X, v(x) < a\}$ so that $\{x : v(x) \in (-\infty, a)\}$, the union of all sets of the form $\{x : x \prec y, v(y) < a\}$, is in $\mathcal{T}$ since each set in the union is in $\mathcal{T}$. Thus $\{x : v(x) \in (-\infty, a)\} \in \mathcal{T}$ for every $a \in \text{Re}$, and, by a symmetric proof, $\{x : v(x) \in (b, \infty)\} \in \mathcal{T}$ for every $b \in \text{Re}$. Since any bounded open interval $(a, b) \in \mathcal{U}$ is the intersection of $(a, \infty) \in \mathcal{U}$ and $(-\infty, b) \in \mathcal{U}$, $\{x : v(x) \in (a, b)\}$ is the intersection of two sets in $\mathcal{T}$ and hence is in $\mathcal{T}$. Since any $A \in \mathcal{U}$ is formed by arbitrary unions and finite intersections of open intervals in Re, the corresponding set $\{x : v(x) \in A\}$ can be formed in a similar way from sets in $\mathcal{T}$ and hence is in $\mathcal{T}$. ◆

Contributions to continuity in the context considered in this subsection have been made also by Eilenberg (1941), Newman and Read (1961), and Rader (1963). Condition 2 of Theorem 3.3 is identical to Condition *B*, p. 160, in Newman and Read. Debreu (1964) includes most of the important results in this area.

Continuity of Increasing Utilities on Re^n

For Re^n we shall let $\mathcal{U}^n$ be the set of all open rectangles along with their arbitrary unions and finite intersections. With X a rectangular subset of Re^n this subsection examines the continuity of u on X with respect to the relative topology $\{A \cap X : A \in \mathcal{U}^n\}$. The following theorem is slightly different than very similar theorems on continuity discussed by Wold (1943), Wold and Jureen (1953), Yokoyama (1956), Debreu (1959), and Newman and Read (1961). The proof is similar to Yokoyama's.

THEOREM 3.6. *The hypotheses of Theorem* 3.3 *imply that there is a real-valued function on X that satisfies* (3.1) *and is continuous in the topology* $\{A \cap X : A \in \mathcal{U}^n\}$.

Proof. Considering Theorems 3.3 and 3.5 we need only show that condition 2 of Theorem 3.5 holds under the stated hypotheses when

$\mathcal{T} = \{A \cap X : A \in \mathcal{U}^n\}$. With $\mathcal{T} = \{A \cap X : A \in \mathcal{U}^n\}$ and $x \prec y$ we show that there is a $T_y \in \mathcal{T}$ such that $y \in T_y$ and $x \prec z$ for every $z \in T_y$. The proof concerning T_x is symmetric to this proof and is left to the reader.

With $x < y$, let $v_i = \inf\{x_i, y_i\}$ and if v_i is greater than some element in X_i let v_i' be any element in X_i less than v_i: otherwise let $v_i' = v_i$. Then $v' \leq v$, $v \leq x$, $v < y$. If $v' = x$, then $x \prec z$ for all $z \neq x, z \in X$, and any T_y containing y but not x suffices. Henceforth we assume that $v' < x$, so that $v' \prec x \prec y$, implying by condition 3 of Theorem 3.3 that for some $\alpha \in (0, 1)$, $x \prec \alpha v' + (1 - \alpha)y$. Now $\alpha v_i' + (1 - \alpha)y_i \leq y_i$ for all i and strict inequality holds for some i. Let $\epsilon > 0$ be smaller than the smallest $y_i - [\alpha v_i' + (1 - \alpha)y_i]$ for which the difference is positive. Then $\alpha v_i' + (1 - \alpha)y_i < y_i - \epsilon$ for all i for which $y_i - [\alpha v_i' + (1 - \alpha)y_i] > 0$. If $v_i' = y_i$, then any z_i less than y_i is not in X_i. Let $T_y' = (y_1 - \epsilon, y_1 + \epsilon) \times (y_2 - \epsilon, y_2 + \epsilon) \times \cdots \times (y_n - \epsilon, y_n + \epsilon)$ and let $T_y = T_y' \cap X$. Then $T_y \in \mathcal{T}$ and for every $z \in T_y$, $\alpha v' + (1 - \alpha)y < z$, so that $x \prec z$ for every $z \in T_y$. ◆

Upper Semicontinuity with Strict Partial Order

Definition 3.4. If $(X, \mathcal{T})$ is a topological space then a real-valued function u on X is *upper semicontinuous* in the topology $\mathcal{T}$ if and only if

$$\{x : x \in X, u(x) < c\} \in \mathcal{T} \qquad \text{for each real number } c. \tag{3.11}$$

Lower semicontinuity is defined by (3.11) after $<$ is changed to $>$. Given a bounded, real-valued function f on X let u on X be defined by

$$u(x) = \inf\{\sup\{f(y) : y \in T\} : x \in T, T \in \mathcal{T}\}. \tag{3.12}$$

For a given real number c suppose $u(x) < c$ for no x. Then $\{x : x \in X, u(x) < c\} = \varnothing$, which is in $\mathcal{T}$. Suppose $u(x) < c$ for some $x \in X$. Then there is a $T_x \in \mathcal{T}$ such that $x \in T_x$ and $\sup\{f(y) : y \in T_x\} < c$. It follows from (3.12) that $u(y) < c$ for every $y \in T_x$. Hence, for each x such that $u(x) < c$ there is a $T_x \in \mathcal{T}$ such that $x \in T_x$ and $u(y) < c$ for every $y \in T_x$. The union of all such T_x will equal $\{x : x \in X, u(x) < c\}$, and this union is in $\mathcal{T}$ by Definition 3.2(3). Hence u is upper semicontinuous in $\mathcal{T}$. We shall use this observation in proving the following theorem.

THEOREM 3.7. *The hypotheses of Theorem* 3.4 *imply that there is a real-valued function on X that satisfies* (3.3) *and is upper semicontinuous in the topology* $\{A \cap X : A \in \mathcal{U}^n\}$.

Proof. As in the proof of Theorem 3.4 let $<^1$ on the non-negative orthant of Re^n be defined as the union of $<$ and $\ll$. From that proof there is a real-valued function f on X that satisfies $x <^1 y \Rightarrow f(x) < f(y)$. By a simple monotonic transformation if necessary, we can suppose that f is bounded.

Then, with u defined as in (3.12), u is upper semicontinuous in the relative topology $\mathcal{T} = \{A \cap X : A \in \mathcal{U}^n\}$.

It remains to show that $x < y \Rightarrow u(x) < u(y)$. Suppose $x < y$. Then, by condition 3 of Theorem 3.4, $z < y$ for some $z \in X$ for which $x \ll z$. There is then an open rectangle $T_x \in \mathcal{T}$ that contains x and has all elements $\ll z$, so that $f(t) < f(z)$ for all $t \in T_x$, so that $u(x) \leq f(z)$. Along with $f(z) < f(y)$ from $z < y$, and $f(y) \leq u(y)$ by the definition of u, this gives $u(x) < u(y)$ as desired. ◆

I am indebted to Hurwicz and Richter (1970) for the approach used in this proof.

3.5 SUMMARY

When X is uncountable and $<$ on X is a weak order, preferences can be faithfully represented by a real-valued function if and only if there is a countable subset Y of X such that whenever $x < y$ there is a $z \in Y$ such that ($x < z$ or $x \sim z$) and ($z \sim y$ or $z < y$). Lexicographic preference orders give examples where this denseness condition fails. With $<$ assumed only to be a strict partial order, we have given a sufficient but not necessary countable order denseness condition for real-valued utilities.

When X is a rectangular subset of n-dimensional Euclidean space and preference increases (or does not decrease) with increases along any dimension, conditions that make better intuitive sense than plain order denseness lead to real-valued utilities.

If (3.1) holds for u on X then there is a continuous (in a specified topology $\mathcal{T}$) utility function on X if and only if $x < y$ implies that there are two subsets of X in $\mathcal{T}$ one of which contains x and has every element less preferred than y and the other of which contains y and has every element preferred to x. The conditions on $<$ used in the weak order and strict partial order theorems for utilities on regions of Re^n also imply the existence of continuous (weak order) and upper semicontinuous (strict partial order) utility functions.

INDEX TO EXERCISES

1. Uncountable sets. 2. Denseness of rationals. 3. Lexicographic order. 4. Theorem 3.1. 5. Theorem 3.2. 6–7. Vector operations. 8. Indifference map. 9–10. Lemma 3.1. 11. Asymmetric transitive closure. 12. Discrete topology. 13. Closed intervals not open. 14. Discontinuity. 15. Theorem 3.5. 16–18. Connected topological spaces. 19. Theorem 3.4. 20. Lower semicontinuity. 21–22. Wold's continuity condition. 23. Convex sets.

Exercises

1. Prove that Re is uncountable by supposing that $\{0\,.\,x_1x_2x_3\cdots: x_i \in \{1, 2\}$ for $i = 1, 2, 3, \ldots\} \subseteq$ Re is countable and showing that this supposition is false. Note also that $\{(x_1, x_2, \ldots): x_i \in \{0, 1\}$ for all $i\}$ is uncountable.

2. Let a and b be numbers with $a < b$. Show that there is a rational number in the open interval (a, b). Use the fact (or axiom) that there is a positive integer n such that $1 < n(b - a)$. Let m be the smallest integer greater than a and show that $m/n \in (a, b)$.

3. For the second example following Theorem 3.1 where $X = [-1, 1]$, show that preferences can be represented by two-dimensional vectors $(u_1(x), u_2(x))$ in Re^2 under a lexicographic order.

4. Prove statement (2) preceding (3.2) in the proof of Theorem 3.1.

5. Describe in your words the effect of E in the proof of Theorem 3.2.

6. Use (3.6) to evaluate: *a.* $(1, 1, 2, 3) + (0, -1, -10, 6)$; *b.* $6(1, 2, 3, 4)$; *c.* $3(0, 0, 1, -1) - (-1, 2, -1, 0)$; *d.* $\alpha(2, 4, -6, -8) + (1 - \alpha)(5, -1, 3, 1)$.

7. The *scalar product* of real vectors $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ is $x \cdot y = x_1y_1 + \cdots + x_ny_n = \sum_{i=1}^{n} x_iy_i$. Evaluate *a.* $(1, 2, 3, 4, 5) \cdot (6, 7, 8, 9, 10)$; *b.* $(3(0, 1, 2) + 4(-2, 1, 3)) \cdot (-5(1, 0, -1))$.

8. Use an indifference map in Re^2 to argue that the hypotheses of Theorem 3.3 do not imply the following: if $x \prec y$ and $0 \leq \alpha < \beta \leq 1$ then $\beta x + (1 - \beta)y \prec \alpha x + (1 - \alpha)y$.

9. Show that (3.7) and (3.8) cannot hold under the stated hypotheses.

10. Show that Lemma 3.1 remains valid when "$x < y < z$" is replaced by "$x \prec y \prec z$ and $x < z$."

11. Prove that the conclusion of Theorem 3.4 remains valid when condition 1 of its hypotheses is replaced by "the transitive closure of $\prec$ on X is asymmetric." (See Exercise 2.5.)

12. The *discrete topology* for any set X is the set of all subsets of X. Is every real-valued function of X continuous in the discrete topology? Why? What does this say about continuity when X is finite?

13. Show that any bounded closed interval $[a, b]$ in Re with $a < b$ is not in $\mathcal{U}$.

14. Let $\prec$ on $X = [0, 2]$ be defined by: $x \prec y$ if ($x < y$ and $x, y \in [0, 1]$) or if ($y < x$ and $x, y \in [1, 2]$); $x \sim (2 - 2x/3)$ when $x \in [0, 1/2)$ and $x \sim (5/3 - 2x/3)$ when $x \in [1/2, 1]$. Show that there is a u on X that satisfies (3.1) and that no such u can be continuous in the relative usual topology.

15. For the proof of Theorem 3.5 show that $\{x: v(x) \in (b, \infty)\} \in \mathcal{T}$ for every $b \in$ Re.

16. A topological space $(X, \mathcal{T})$ is *connected* if X cannot be partitioned into two nonempty subsets both of which are in $\mathcal{T}$. Prove that if $(X, \mathcal{T})$ is connected, if u

on X is continuous in $\mathcal{T}$, and if $u(x) < u(y)$ for $x, y \in X$, then for each $c \in (u(x), u(y))$ there is a $z \in X$ such that $u(z) = c$.

17. Show that any rectangular subset of Re^n is connected.

18. Let X be a rectangular subset of Re^n. With $x, y \in X$, the line segment $L = \{\alpha x + (1 - \alpha)y : \alpha \in [0, 1]\}$ between x and y has the relative topology $\mathcal{T}' = \{A \cap L : A \in \mathcal{U}^n\}$. (*a*) Given the result of the preceding exercise show that $(L, \mathcal{T}')$ is connected. (*b*) Suppose u on X is continuous in $\{A \cap X : A \in \mathcal{U}^n\}$, and let $u'(z) = u(z)$ when $z \in L$. Argue that u' on L is continuous in $\mathcal{T}'$.

19. In the proof of Theorem 3.4 show that if $x \prec y$, then there is a $T_x \in \mathcal{T} = \{A \cap X : A \in \mathcal{U}^n\}$ such that $x \in T_x$ and $z \prec y$ for every $z \in T_x$.

20. Let f be a bounded, real-valued function on X and let v on X be defined by $v(x) = \sup\{\inf\{f(y) : y \in T\} : x \in T, T \in \mathcal{T}\}$. Show that v is lower semicontinuous in the topology $\mathcal{T}$.

21. Wold (1943). Condition W: *if* $x \prec y$ *and* $y \prec z$ *then* $\alpha x + (1 - \alpha) z \sim y$ *for some* $\alpha \in (0, 1)$. Show that the conclusions of Theorems 3.3 and 3.6 remain valid when condition W replaces condition 3 of Theorem 3.3. Also show by indifference curves in Re^2 that α need not be unique. (See Exercise 8.)

22. Use the results of Exercises 16 and 18 to show that if X is a rectangular subset in Re^n, if u on X is continuous in $\{A \cap X : A \in \mathcal{U}^n\}$, and if conditions 1 and 2 of Theorem 3.3 hold, then condition 3 and condition W (Exercise 21) must hold also.

23. $X \subseteq \mathrm{Re}^n$ is *convex* $\Leftrightarrow \alpha x + (1 - \alpha)y \in X$ whenever $x, y \in X$ and $\alpha \in (0, 1)$. Show that "X is convex" and "$(X, \{A \cap X : A \in \mathcal{U}^n\})$ is not connected" cannot both be true. Assuming that X is convex, use this result along with that of Exercise 16 to conclude that if there is a real-valued function u on X that satisfies (3.1) and is continuous in $\{A \cap X : A \in \mathcal{U}^n\}$ then condition 3 of Theorem 3.3 and condition W must be true. Thus, regardless of whether condition 2 of Theorem 3.3 holds, condition 3 must hold when X is a convex subset of Re^n in order that there be a u on X that satisfies (3.1) and is continuous. But note also from Exercise 14 that there can be a u on X satisfying (3.1) when condition 3 fails and X is convex.

Chapter 4

ADDITIVE UTILITIES WITH FINITE SETS

Except for Chapter 6, the remaining chapters of Part I examine special kinds of preferences and utilities that might arise in multiple-factor situations. Chapter 3 has already considered some basic theory for n-dimensional Euclidean spaces. This chapter and the next deal with additive utility representations for preference orders on sets of n-tuples. Section 4.3 considers lexicographic utility.

Throughout this chapter we shall usually assume that X is a nonempty subset of the Cartesian product

$$\prod_{i=1}^{n} X_i = X_1 \times X_2 \times \cdots \times X_n$$

of n other finite sets. Thus, each alternative in X is an n-tuple $x = (x_1, \ldots, x_n)$. Each X_i is a factor or attribute set. For convenience we assume that each $x_i \in X_i$ is the ith component of some $x \in X$.

The subscript i could refer to n different attributes or performance characteristics of competing alternatives, it could refer to a time factor (n periods), and so forth. We shall identify conditions for $\prec$ on X that lead to additive utility representations such as the one for weak orders: $x \prec y \Leftrightarrow u_1(x_1) + \cdots + u_n(x_n) < u_1(y_1) + \cdots + u_n(y_n)$.

It should be emphasized that $\prec$ is applied to pairs of complete n-tuples, or whole alternatives. In multiple-factor situations it often seems natural to think in terms of a preference order for each factor and then to wonder how these ought to be combined or synthesized into an overall preference order. However, this approach presupposes a certain kind of independence among the factors, namely that the order for a given factor is independent of the particular levels of the other factors. This can of course be false. For example, suppose that (chicken for dinner tonight, chicken for dinner tomorrow night) $\prec$ (steak tonight, steak tomorrow night) $\prec$ (chicken tonight, steak

tomorrow night) $\prec$ (steak tonight, chicken tomorrow night). In this case, preference for tonight clearly depends on what is assumed about tomorrow night. Under the hypothesis of chicken tomorrow, steak is preferred tonight. Under the hypothesis of steak tomorrow, chicken is preferred tonight.

For situations where the independence conditions seem reasonable and additive utilities apply, Fishburn (1967) summarizes a number of ways to estimate factor utilities so as to satisfy the additive representation.

4.1 PREFERENCE INDEPENDENCE AMONG FACTORS

Consider a two-dimensional case where $X = X_1 \times X_2$, $\prec$ is a weak order and, for each $x_1, y_1 \in X_1$ and $x_2, y_2 \in X_2$,

$$(x_1, x_2) \prec (y_1, x_2) \Rightarrow (x_1, y_2) \prec (y_1, y_2), \tag{4.1}$$

$$(x_1, x_2) \prec (x_1, y_2) \Rightarrow (y_1, x_2) \prec (y_1, y_2). \tag{4.2}$$

The first of these says that, if we define $x_1 \prec_1 y_1 \Leftrightarrow (x_1, x_2) \prec (y_1, x_2)$ for some $x_2 \in X_2$, then $\prec_1$ is a weak order on X_1 that is independent of the particular element used from X_2. Similarly, the second says that, when the first factor is fixed, there will be a weak order $\prec_2$ on X_2 derived in the natural way from $\prec$ that does not depend on the element used from X_1. In the simplest possible way this suggests that X_1 and X_2 are independent in a preference sense.

As demonstrated by Scott and Suppes (1958), even in the two-dimensional case considered above it may be necessary to go beyond (4.1) and (4.2) to obtain an additive-utility representation of the form $(x_1, x_2) \prec (y_1, y_2) \Leftrightarrow u_1(x_1) + u_2(x_2) < u_1(y_1) + u_2(y_2)$. Clearly, (4.1) and (4.2) are necessary for the existence of such a representation, but they are not sufficient. Suppose for example that $\prec$ on $X = \{1, 2, 3\} \times \{1, 3, 5\}$ is a weak order with

$$(x_1, x_2) \prec (y_1, y_2) \Leftrightarrow x_1x_2 + (x_1)^{x_2} < y_1y_2 + (y_1)^{y_2}. \tag{4.3}$$

Since $u(x_1, x_2) = x_1x_2 + (x_1)^{x_2}$ is strictly increasing in x_1 for any fixed x_2 and is strictly increasing in x_2 for any fixed x_1, (4.1) and (4.2) hold. However, additive utilities do not exist. To the contrary, suppose that there are real-valued functions u_1 on $X_1 = \{1, 2, 3\}$ and u_2 on $X_2 = \{1, 3, 5\}$ such that $(x_1, x_2) \prec (y_1, y_2) \Leftrightarrow u_1(x_1) + u_2(x_2) < u_1(y_1) + u_2(y_2)$. Then, since $(2, 1) \sim (1, 3)$ and $(1, 5) \sim (3, 1)$ by (4.3),

$$u_1(2) + u_2(1) = u_1(1) + u_2(3)$$

$$u_1(1) + u_2(5) = u_1(3) + u_2(1).$$

By adding these equalities and cancelling identical terms we get

$$u_1(2) + u_2(5) = u_1(3) + u_2(3)$$

which, according to the presumed existence of an additive representation, yields $(2, 5) \sim (3, 3)$. But, by (4.3), $u(2, 5) = 42$ and $u(3, 3) = 36$, so that $(3, 3) < (2, 5)$. Hence there is no additive representation for this case.

Additive Utilities

In generalizing independence conditions like (4.1) and (4.2) we shall use a sequence of equivalence relations E_m on X^m ($m = 2, 3, \ldots$) where X^m is the m-fold Cartesian product of X with itself.

Definition 4.1. $(x^1, \ldots, x^m)\, E_m\, (y^1, \ldots, y^m)$ if and only if $m > 1$, x^j, $y^j \in X$ for $j = 1, \ldots, m$, and with $X \subseteq \prod_{i=1}^n X_i$ it is true for each i that $x_i^1, \ldots, x_i^m$ is a permutation (reordering) of $y_i^1, \ldots, y_i^m$.

Thus, for (4.1), $((x_1, x_2), (y_1, y_2))\, E_2\, ((y_1, x_2), (x_1, y_2))$, and in the example refuting additivity for (4.3), $((2, 1), (1, 5), (3, 3))\, E_3\, ((1, 3), (3, 1), (2, 5))$. With $n = 3$ and $(x_1, x_2, x_3) =$ (net profit, market share, dividend per share of stock), the following arrays reveal that $(x^1, \ldots, x^4)\, E_4\, (y^1, \ldots, y^4)$.

	profit	share	dividend
x^1	\$1m	20%	30¢
x^2	\$0m	10%	50¢
x^3	\$2m	30%	45¢
x^4	−\$1m	15%	10¢

	profit	share	dividend
y^1	\$2m	20%	50¢
y^2	−\$1m	10%	45¢
y^3	\$1m	15%	10¢
y^4	\$0m	30%	30¢

For the purpose of further discussion we shall first present three additive utility theorems. For comparative convenience they are presented together in Theorem 4.1. There is a theorem A, a theorem B, and a theorem C, with "hypotheses" and "conclusions" noted accordingly.

THEOREM 4.1. *Suppose $X \subseteq \prod_{i=1}^n X_i$ is finite. Then*

A. $[(x^1, \ldots, x^m)\, E_m\, (y^1, \ldots, y^m)$, $x^j < y^j$ *or* $x^j = y^j$ *for* $j = 1, \ldots, m - 1] \Rightarrow$ *not* $x^m < y^m$;

B. $[(x^1, \ldots, x^m)\, E_m\, (y^1, \ldots, y^m)$, $x^j < y^j$ *or* $x^j \approx y^j$ *for* $j = 1, \ldots, m - 1] \Rightarrow$ *not* $x^m < y^m$;

C. $[(x^1, \ldots, x^m)\, E_m\, (y^1, \ldots, y^m)$, $x^j < y^j$ *or* $x^j \sim y^j$ *for* $j = 1, \ldots, m - 1] \Leftrightarrow$ *not* $x^m < y^m$;

for all $x^1, \ldots, x^m, y^1, \ldots, y^m \in X$ *and* $m = 2, 3, \ldots$, *if and only if there are real-valued functions* $u_1, \ldots, u_n$ *on* $X_1, \ldots, X_n$ *respectively such that, for all* $x, y \in X$,

A^*. $x < y \Rightarrow \sum_{i=1}^n u_i(x_i) < \sum_{i=1}^n u_i(y_i)$;

B^*. $x < y \Rightarrow \sum_{i=1}^n u_i(x_i) < \sum_{i=1}^n u_i(y_i), x \approx y \Rightarrow \sum_{i=1}^n u_i(x_i) = \sum_{i=1}^n u_i(y_i)$;

C^*. $x < y \Leftrightarrow \sum_{i=1}^n u_i(x_i) < \sum_{i=1}^n u_i(y_i)$.

Indifference ($\sim$) and $\approx$ are defined as $x \sim y \Leftrightarrow$ (not $x < y$, not $y < x$) and $x \approx y \Leftrightarrow (z \sim x \Leftrightarrow z \sim y$, for all $z \in X)$, as in Chapters 2 and 3.

Unlike (4.1) and (4.2), the conclusions of A, B, and C are stated in the negative. It is easily seen that A is necessary for A^*, that B is necessary for B^*, and that C is necessary for C^*. For example, suppose that A^* holds and that the hypotheses of A hold with $(x^1, \ldots, x^m)\, E_m\, (y^1, \ldots, y^m)$ and $x^j < y^j$ or $x^j = y^j$ for each $j < m$. Then, by A^*, $\sum_{j=1}^{m-1} \sum_{i=1}^{n} u_i(x_i^j) \leq \sum_{j=1}^{m-1} \sum_{i=1}^{n} u_i(y_i^j)$. But, by E_m, $\sum_{j=1}^{m} \sum_{i=1}^{n} u_i(x_i^j) = \sum_{j=1}^{m} \sum_{i=1}^{n} u_i(y_i^j)$. Therefore $\sum_{i=1}^{n} u_i(y_i^m) \leq \sum_{i=1}^{n} u_i(x_i^m)$, which by A^* implies not $x^m < y^m$, which is the conclusion of A.

We shall consider the sufficiency of A for A^*, B for B^*, and C for C^* in the next section. These sufficiency proofs will be based on a theorem from linear algebra called the Theorem of The Alternative, which will be proved in the next section.

Further Remarks on Independence Conditions

Each of conditions A, B, and C in Theorem 4.1 is actually a denumerable bundle of conditions, one for each equivalence E_m, $m = 2, 3, \ldots$. If we let A_m, B_m, and C_m denote the part of condition A, B, and C that applies to E_m, then $A_{m+1} \Rightarrow A_m$, $B_{m+1} \Rightarrow B_m$, and $C_{m+1} \Rightarrow C_m$ for all $m \geq 2$. However, as suggested by Scott and Suppes (1958), there is no one finite value of m for which $A_m \Rightarrow A^*$ or $B_m \Rightarrow B^*$ or $C_m \Rightarrow C^*$ for all finite sets X. We now consider some of the other aspects of A, B, and C.

Our main purpose in including $x^j = y^j$ in the hypotheses of A was to get $A_{m+1} \Rightarrow A_m$, but the equality part of the hypothesis of A is unnecessary. Although A does not imply that $<$ is a strict partial order since it does not imply transitivity, it does say that if $x^1 < x^2, x^2 < x^3, \ldots, x^{m-1} < x^m$ then not $x^m < x^1$. This follows from the fact that $(x^1, x^2, \ldots, x^m)\, E_m\, (x^2, \ldots, x^m, x^1)$. Hence when A holds, the transitive closure of $<$ (Exercise 2.5) is a strict partial order.

Like A, B does not imply that $<$ is a strict partial order. For example, suppose $X = \{x, y, z, t\}$ and $< \; = \{x < y, y < z, x < t\}$ with $\sim$ elsewhere. Then $\approx$ holds for no distinct pair of elements in X so that B reduces, in effect, to A. Since A is consistent with $<$ as given and $<$ is not transitive, B does not imply that $<$ is a strict partial order.

On the other hand, B does imply that $\approx$ is an equivalence since it implies asymmetry on considering $(x, y)\, E_2\, (y, x)$, and asymmetry of $<$ implies that $\approx$ is an equivalence (Exercise 2.3). B implies also, as in the conclusions of Theorem 2.3, that $(x < y, y \approx z) \Rightarrow x < z$ and that $(x \approx y, y < z) \Rightarrow x < z$. For example, since $(x, y, z)\, E_3\, (y, z, x)$, $x < y$ and $y \approx z$ imply not $z < x$ by B. Hence either $x < z$ or $x \sim z$. If $x \sim z$ then $x \sim y$ by the definition for $y \approx z$. But $x < y$. Hence $x < z$.

C of course implies that $\prec$ is a weak order. Suppose not $x \prec y$ and not $y \prec z$. Then $y \prec x$ or $y \sim x$, and $z \prec y$ or $z \sim y$, so that, since $(y, z, x)\, E_3\, (x, y, z)$, C implies not $x \prec z$. Hence, C implies that $\prec$ is negatively transitive. Asymmetry follows from $(x, y)\, E_2\, (y, x)$.

Remarks on Additive Utilities

It should be noted that if additive utilities exist in the sense of A^*, B^*, or C^*, then it does not follow that any utility function u on X that preserves the preference order can be written in an additive form. For example suppose in connection with C^* that $x \prec y \Leftrightarrow u(x) < u(y)$. Then it may be impossible to write u in an additive form when C^* holds. What C^* says is that, among all functions u that satisfy $x \prec y \Leftrightarrow u(x) < u(y)$, there is at least one that can be written in the additive form as $u(x) = u_1(x_1) + \cdots + u_n(x_n)$.

It cannot be emphasized too strongly that additive utilities might not exist in some situations where their use seems attractive for ease in analysis. Possibly the best way to test condition A, or B, or C, is to try deliberately to find n-tuples in X that violate the condition. An inability to construct a violation would lend support to the credibility of the condition. Another obvious way of testing for additivity is to obtain a set of preference statements, convert these into additive utility inequalities and equalities (for C when $\sim$ arises) and test this system for the existence of a solution. If no solution exists then a violation of the appropriate condition has been uncovered.

4.2 THEOREM OF THE ALTERNATIVE

To prove the sufficiency of the conditions A, B, and C of Theorem 4.1, we shall use the following theorem, which is discussed by Tucker (1956, p. 10), Goldman (1956), and Aumann (1964, p. 225), and which has been used by Tversky (1964), Scott (1964), and Adams (1965) to prove theorems like Theorem 4.1. Re^N is N-dimensional Euclidean space and $c \cdot x^k = \sum_{j=1}^N c_j x_j^k$.

THEOREM 4.2 (THEOREM OF THE ALTERNATIVE). *If $x^1, \ldots, x^M \in \mathrm{Re}^N$ and $1 \le K \le M$, then either there is a $c \in \mathrm{Re}^N$ such that*

$$c \cdot x^k > 0 \quad \text{for} \quad k = 1, \ldots, K \tag{4.4}$$

$$c \cdot x^k = 0 \quad \text{for} \quad k = K + 1, \ldots, M, \tag{4.5}$$

or there are non-negative numbers $r_1, \ldots, r_K$ not all of which equal zero and numbers $r_{K+1}, \ldots, r_M$ such that

$$\sum_{k=1}^{M} r_k x_j^k = 0 \quad \textit{for} \quad j = 1, \ldots, N. \tag{4.6}$$

Proof. Let $S = \{x^1, \ldots, x^K\}$ and $T = \{x^{K+1}, \ldots, x^M\}$. Let $\bar{S}$ be the convex closure of S so that

$$\bar{S} = \left\{x : x = \sum_{i=1}^{m} \lambda_i a^i \text{ with } \sum \lambda_i = 1,\ m > 0, \text{ and } \lambda_i \geq 0 \text{ and } a^i \in S \text{ for all } i\right\},$$

and let T' be the vector space generated by T so that

$$T' = \left\{x : x = \sum_{i=1}^{m} \sigma_i b^i \text{ with } m > 0 \text{ and } \sigma_i \in \text{Re and } b^i \in T \text{ for all } i\right\} \cup \mathbf{0}$$

where $\mathbf{0}$ is the origin of Re^N. When $K = M$, $T = \varnothing$ and $T' = \mathbf{0}$.

The two alternatives depend on whether $\bar{S}$ and T' have a common element. If $\bar{S} \cap T' \neq \varnothing$ then (4.6) holds as is seen from $\sum \lambda_i a^i = \mathbf{0}$ or $\sum \lambda_i a^i - \sum \sigma_i b^i = \mathbf{0}$ with the obvious definitions of the r_k in terms of the λ_i $(k \leq K)$ and σ_i $(k > K)$.

On the other hand, (4.4)–(4.5) hold when $\bar{S} \cap T' = \varnothing$. Since both S and T are finite sets (and this is critical to the conclusion), it can be shown that there are vectors $s \in \bar{S}$ and $t \in T'$ such that $(x - y)^2 \geq (s - t)^2 > 0$ when $x \in \bar{S}$ and $y \in T'$. The $x^2 = x \cdot x$ (not to be confused with $x^2 \in S$). Let $x \in \bar{S}$. Then, with $0 \leq \lambda \leq 1$, $(1 - \lambda)s + \lambda x \in \bar{S}$. Since $t \in T'$, $(1 - \lambda)t \in T'$. Hence $[(1 - \lambda)s + \lambda x - (1 - \lambda)t]^2 \geq (s - t)^2$, which reduces to $2\lambda(s - t) \cdot (x - (s - t)) + \lambda^2((s - t) - x)^2 \geq 0$. Take $\lambda > 0$, divide by λ, and let λ approach 0: this leaves $(s - t) \cdot (x - (s - t)) \geq 0$, or $(s - t) \cdot x \geq (s - t)^2 > 0$, or $(s - t) \cdot x > 0$. Let $c = s - t$. Thus $c \cdot x > 0$ for all $x \in \bar{S}$, so that (4.4) holds.

To verify (4.5) when $K < M$ and $\bar{S} \cap T' = \varnothing$, take $y \in T'$. Then $\sigma y + t \in T'$ so that $(\sigma y + t - s)^2 \geq (s - t)^2$, or $\sigma^2 y^2 \geq 2\sigma y \cdot (s - t) = 2\sigma c \cdot y$. First, take $\sigma > 0$, divide by σ and let σ approach 0. Since $y^2 \geq 0$ this leaves $0 \geq c \cdot y$. Second, take $\sigma < 0$ and divide by σ giving $\sigma y^2 \leq 2c \cdot y$. Letting σ approach 0 from below gives $0 \leq c \cdot y$. Hence $c \cdot y = 0$. ◆

In the following I shall detail only the proof that $B \Rightarrow B^*$ in Theorem 4.1. The proof that $C \Rightarrow C^*$ is entirely similar since C^* is equivalent to $x < y \Rightarrow \sum u_i(x_i) < \sum u_i(y_i)$ and $x \sim y \Rightarrow \sum u_i(x_i) = \sum u_i(y_i)$, which is like B^* with $\approx$ replaced by $\sim$. The proof that $A \Rightarrow A^*$, as given by Adams (1965), involves only (4.4) from Theorem 4.2 and not (4.5) since there are no equality implications in A^*.

Sufficiency Proof of Theorem 4.1B. Let B hold. For the application of Theorem 4.2 we let N equal the size of X_1 plus the size of $X_2 \cdots$ plus the size of X_n and let $c = (u_1(x_{11}), u_1(x_{12}), \ldots, u_n(x_{nt}))$ with N components. Let K be the size of $<$ (the number of $x < y$ statements) and let $M - K$ be half the size of $\approx - =$, containing exactly one of $x \approx y$ and $y \approx x$ for each such

x, y pair for which $x \not\approx y$. The $K <$ statements in the conclusion B^* and the $M - K =$ statements translate into the equivalent system

$$c \cdot a^k > 0 \quad \text{for} \quad k = 1, \ldots, K \; (x^k < y^k) \tag{4.7}$$

$$c \cdot a^k = 0 \quad \text{for} \quad k = K + 1, \ldots, M \; (x^k \approx y^k) \tag{4.8}$$

where each $a_j^k \in \{-1, 0, 1\}$ and $\sum_{j=1}^{N} a_j^k = 0$ for each k. B^* holds if and only if (4.7) and (4.8) have a c solution.

Suppose there is no such c solution. Then, by Theorem 4.2, there are $r_k \geq 0$ for $k = 1, \ldots, K$ with $r_k > 0$ for some $k \leq K$, and $r_{K+1}, \ldots, r_M$ such that

$$\sum_{k=1}^{M} r_k a_j^k = 0 \quad \text{for} \quad j = 1, \ldots, N. \tag{4.9}$$

Because each a_j^k is rational there is a set of rational and hence integer r_k that satisfy (4.9). If some of these integer r_k for $k > K$ are negative they can be made positive by replacing a^k with $-a^k$ in (4.8) and (4.9) and replacing r_k by $-r_k$ in (4.9), which does not essentially change (4.8) or (4.9) and is legitimate from the standpoint of (4.8) since $\approx$ is symmetric. Then, with all $r_k \geq 0$, (4.9) says that $(r_1 x^1\text{'s}, r_2 x^2\text{'s}, \ldots, r_M x^M\text{'s}) \; E_{r_1+r_2+\cdots+r_M} \; (r_1 y^1\text{'s}, r_2 y^2\text{'s}, \ldots, r_M y^M\text{'s})$ with $x^k < y^k$ for $k = 1, \ldots, K$ and $x^k \approx y^k$ for $k = K + 1, \ldots, M$. Since some $r_k > 0$ for $k \leq K$ it follows that B does not hold, for if $\sum r_k = 1$ then irreflexivity of $<$ (implied by B) is violated and if $\sum r_k > 1$ then B is violated as it stands. Since B is in fact assumed to hold it must be false that there is no c solution for (4.7) and (4.8). ◆

4.3 LEXICOGRAPHIC UTILITIES

The purpose of this section is to note an affinity between additive utilities and lexicographic utilities. For the latter case we define $<^L$ for real vectors $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$:

$$a <^L b \Leftrightarrow a \neq b \quad \text{and} \quad b_k < a_k \Rightarrow a_j < b_j \quad \text{for some } j < k, \qquad k = 2, \ldots, n. \tag{4.10}$$

Thus $a <^L b \Leftrightarrow a_1 < b_1$ or $[a_1 = b_1, a_2 < b_2]$ or $\cdots$ or $[a_1 = b_1, \ldots, a_{n-1} = b_{n-1}, a_n < b_n]$.

In comparison with Theorem 4.1A we shall consider the existence of real-valued functions $u_1, \ldots, u_n$ on $X_1, \ldots, X_n$ such that

$$x < y \Rightarrow (u_1(x_1), \ldots, u_n(x_n)) <^L (u_1(y_1), \ldots, u_n(y_n)). \tag{4.11}$$

The comparison for Theorem 4.1C is

$$x < y \Leftrightarrow (u_1(x_1), \ldots, u_n(x_n)) <^L (u_1(y_1), \ldots, u_n(y_n)). \tag{4.12}$$

In both cases the order of the X_i is very significant. In a preference sense, (4.11) or (4.12) says approximately that X_1 dominates X_2, X_2 dominates X_3, and so forth.

The main point to be made about (4.11) and (4.12) is that condition A of Theorem 4.1 is necessary for (4.11), and condition C is necessary for (4.12). For example, suppose (4.11) holds along with the hypotheses of condition A: $(x^1, \ldots, x^m)\ E_m\ (y^1, \ldots, y^m)$ and $x^j < y^j$ or $x^j = y^j$ for $j = 1, \ldots, m-1$. Then $u_1(x_1^j) \le u_1(y_1^j)$ for all $j < m$, and since $\sum_1^m u_1(x_1^j) = \sum_1^m u_1(y_1^j)$ by E_m, $u_1(y_1^m) \le u_1(x_1^m)$. If $u_1(x_1^j) < u_1(y_1^j)$ for some $j < m$ then $u_1(y_1^m) < u_1(x_1^m)$ so that $(u_1(y_1^m), \ldots, u_n(y_n^m)) <^L (u_1(x_1^m), \ldots, u_n(x_n^m))$. If $u_1(x_1^j) = u_1(y_1^j)$ for all $j < m$ then $u_1(y_1^m) = u_1(x_1^m)$, in which case we repeat the analysis just given, using u_2 instead of u_1. Continuing this we conclude that either $(u_1(y_1^m), \ldots, u_n(y_n^m)) <^L (u_1(x_1^m), \ldots, u_n(x_n^m))$ or else that the two utility vectors are equal. It then follows from (4.11) that not $x^m < y^m$, which is the conclusion of condition A.

Thus, if X is finite and lexicographic utilities exist in the sense of (4.11) then additive utilities exist in the sense of A^* in Theorem 4.1. A similar assertion holds for (4.12) and C^*. Clearly, the converses of these assertions are not generally valid.

What is required for (4.12) in addition to condition C? Clearly, something like the following is needed.

Condition L. *If $x < y$ when $x_i = y_i$ for all i except $i = k$, then $x^* < y^*$ when $(x_i^* = x_i, y_i^* = y_i)$ for all $i \le k$, provided $x, y, x^*, y^* \in X$.*

Interestingly enough, when this lexicographic dominance condition is used and $X = \prod X_i$, it is no longer necessary to use all of condition C. The following uses only the $m = 2$ part of C. We can also remove the strict finiteness assumption for X.

THEOREM 4.3. *Suppose X is countable and $X = \prod_{i=1}^n X_i$. Then* (4.12) *holds for all $x, y \in X$ if and only if*

1. *$\prec$ is negatively transitive,*
2. $[(x, z)\ E_2\ (y, w), x < y$ *or* $x \sim y] \Rightarrow$ *not* $z < w$,
3. *condition L holds.*

Sufficiency Proof. Under the given conditions define $x_i <_i y_i \Leftrightarrow z < w$ for some $z, w \in X$ such that $z_j = w_j$ for all $j \ne i$ and $(z_i = x_i, w_i = y_i)$. Then $<_i$ on X_i is a weak order: asymmetry follows from condition 2 and negative transitivity follows from condition 1 and $X = \prod X_i$. Then, by Theorem 2.2, there is, for each i, a real-valued function u_i on X_i such that $x_i <_i y_i \Leftrightarrow u_i(x_i) < u_i(y_i)$.

Suppose that $(u_1(x_1), \ldots, u_n(x_n)) <^L (u_1(y_1), \ldots, u_n(y_n))$ and let t be the smallest i for which $u_i(x_i) < u_i(y_i)$, with $u_i(x_i) = u_i(y_i)$ for all $i < t$. We wish

to show that $x < y$. To do this we first note that, if $1 < t$, (not $x_1 <_1 y_1$, not $y_1 <_1 x_1$) $\Rightarrow (x_1, x_2, \ldots, x_n) \sim (y_1, x_2, \ldots, x_n)$. Similarly, if $2 < t$, $u_2(x_2) = u_2(y_2) \Rightarrow (y_1, x_2, \ldots, x_n) \sim (y_1, y_2, x_3, \ldots, x_n)$. Continuing this and using the transitivity of $\sim$ (from weak order), we get $(x_1, \ldots, x_n) \sim (y_1, \ldots, y_{t-1}, x_t, \ldots, x_n)$. Now $x_t <_t y_t$. Therefore $(y_1, \ldots, y_{t-1}, x_t, \ldots, x_n) < (y_1, \ldots, y_{t-1}, y_t, x_{t+1}, \ldots, x_n)$ by the definition of $<_t$ and condition 2. Hence, by condition L, $(y_1, \ldots, y_{t-1}, x_t, \ldots, x_n) < (y_1, \ldots, y_{t-1}, y_t, y_{t+1}, \ldots, y_n)$. Thus, by Theorem 2.1, $x < y$.

On the other hand suppose that not $(u_1(x_1), \ldots, u_n(x_n)) <^L (u_1(y_1), \ldots, u_n(y_n))$. Then either $(u_1(y_1), \ldots, u_n(y_n)) <^L (u_1(x_1), \ldots, u_n(x_n))$, in which case $y < x$ and hence not $x < y$, or else the two utility vectors are equal, in which case the $\sim$ analysis of the preceding paragraph leads to $x \sim y$ and hence not $x < y$. ◆

4.4 SUMMARY

When X is a finite subset of the Cartesian product of n other sets, additive utilities for several cases considered exist if and only if appropriate independence conditions hold. The finiteness of X is crucial for these cases in the absence of additional conditions. However, in a weak order case with $X = \prod X_i$, the finiteness condition can be replaced by a countability condition in a simple axiomatization for lexicographic utility according to a definite dominance order for the n factors. With X finite, lexicographic utilities imply the existence of additive utilities, but the converse is not generally true.

It cannot be emphasized too strongly that additive utilities might not exist in some situations where their use seems attractive for ease in analysis.

INDEX TO EXERCISES

1. The size problem. 2. Weak orders and additivity. 3–4. Functional forms which may or may not admit additive utilities. 5. Condition C_4. 6. The necessity of all of C. 7–8. Variations on C. 9. E_m. 10. Necessity of B^* and C^*. 11–12. B and A. 13–18. Applications of the Theorem of The Alternative. 18. Proof of Theorem 2.9. 19. $<^L$. 20. Admissible transformations.

Exercises

1. Suppose $X = \prod_{i=1}^{10} X_i$ and each X_i has 10 elements. Then X has 10 billion elements but there are only 100 x_i. Discuss the potential attractiveness of additive

utilities from the standpoint of size and the number of utility values that need to be estimated.

2. Let $X = \{a, b\} \times \{c, d\}$. How many preference weak orders can be defined on X? List those for which additive utilities exist as in Theorem 4.1C^*.

3. With $X = X_1 \times X_2$, $X_i = \{1, 2, \ldots, M\}$ with M large, and $x \prec y \Leftrightarrow u(x_1, x_2) < u(y_1, y_2)$, for which of the following cases do there exist u_1 and u_2 so that $x \prec y \Leftrightarrow u_1(x_1) + u_2(x_2) < u_1(y_1) + u_2(y_2)$? (*a*) $u(x_1, x_2) = x_1x_2$; (*b*) $u(x_1, x_2) = x_1^2 + x_1x_2$; (*c*) $u(x_1, x_2) = x_1 + x_2 + x_1x_2$; (*d*) $u(x_1, x_2) = \sup\{x_1, x_2\}$; (*e*) $u(x_1, x_2) = |x_1 - x_2|$ (absolute value); (*f*) $u(x_1, x_2) = 1/(x_1x_2)$; (*g*) $u(x_1, x_2) = x_1/(x_1 + x_2)$. For each that admits additive utilities, tell why this is so.

4. Let $X = X_1 \times X_2$, with X_1 and X_2 sets of positive integers and suppose that $u(x_1, x_2) = x_1x_2 + (x_1x_2)^3$ and $x \prec y \Leftrightarrow u(x_1, x_2) < u(y_1, y_2)$. Show that additive utilities exist in the sense of Theorem 4.1C^*.

5. The accompanying utility matrix gives $u(a, p)$ for $(a, p) \in X = \{a_1, \ldots, a_4\} \times \{p_1, \ldots, p_4\}$. Assume that $(a, p) \prec (a', p') \Leftrightarrow u(a, p) < u(a', p')$,

	p_1	p_2	p_3	p_4
a_1	0	4	8	9
a_2	5	9	12	14
a_3	8	11	13	15
a_4	10	15	16	17

and show that condition C_4 (C with $m = 4$) of Theorem 4.1 fails. Does condition C_3 hold?

6. Let $X = \{(1, 1), (2, 2), \ldots, (m, m), (1, 2), (2, 3), \ldots, (m - 1, m), (m, 1)\}$. Let $u(j, k) = 0$ for all $(j, k) \in X$ except for $(m, 1)$ where $u(m, 1) > 0$, and take $x \prec y \Leftrightarrow u(x) < u(y)$. Show that condition C_m of Theorem 4.1 fails but that condition C_{m-1} holds.

7. Show that condition C of Theorem 4.1 implies that if $(x^1, \ldots, x^m)E_m(y^1, \ldots, y^m)$ and $x^j \prec y^j$ for all $j < m$ then $y^m \prec x^m$.

8. (*Continuation.*) Tversky (1964) uses an axiom he calls the *cancellation law*, which in our terms reads: *if* $(x^1, \ldots, x^m)\, E_m\, (y^1, \ldots, y^m)$, *if* $x^j \prec y^j$ *or* $x^j \sim y^j$ *for all* $j < m$, *and if* $x^j \prec y^j$ *for some* $j < m$, *then* $y^m \prec x^m$. With $\sim$ defined by (2.2) as usual, show that Tversky's condition is implied by C and that C is implied by Tversky's condition plus the assumption that $\prec$ is irreflexive.

9. Prove that E_m on X^m is an equivalence.

10. Show that $B^* \Rightarrow B$ and that $C^* \Rightarrow C$.

11. Prove that B implies that $x \prec z$ when $x \approx y$ and $y \prec z$.

12. Does A of Theorem 4.1 imply that $\approx$ is an equivalence? Why?

13. Write out the sufficiency proof for Theorem 4.1A.

14. Write out the sufficiency proof for Theorem 4.1C.

15. With $<^1$ defined by (2.12) for interval orders, let condition D be: $[(x^1, \ldots, x^m)\ E_m\ (y^1, \ldots, y^m),\ x^j <^1 y^j$ *for* $j = 1, \ldots, m-1] \Rightarrow$ *not* $x^m <^1 y^m$. Suppose $X \subseteq \prod X_i$ is finite. Prove that $\prec$ on X is an interval order and condition D holds if and only if there are real-valued functions $u_1, \ldots, u_n$ on $X_1, \ldots, X_n$ respectively and a nonnegative real-valued function σ on X such that, for all $x, y \in X$,

$$x \prec y \Leftrightarrow \sum_{i=1}^{n} u_i(x_i) + \sigma(x) < \sum_{i=1}^{n} u_i(y_i).$$

Use the Theorem of The Alternative in your sufficiency proof.

16. Let condition E be: $[(x^1, \ldots, x^{2m})\ E_m\ (y^1, \ldots, y^{2m}),\ x^j \sim y^j$ *for* $j = 1, \ldots, m$ *and* $x^j \prec y^j$ *for* $j = m+1, \ldots, 2m-1] \Rightarrow$ *not* $x^{2m} \prec y^{2m}$. Show that if E holds and not $x \prec x$ for some $x \in X$, then $\prec$ on X is irreflexive and asymmetric, $\prec$ is transitive, and $\prec$ satisfies $p10$ and $p11$ of Section 2.4 and hence is a semiorder. (Do not use the Theorem of The Alternative here.) Note the necessity of using not $x \prec x$ for some $x \in X$, for without this we could have $X = \{x\}$ and $x \prec x$ with condition E holding.

17. (*Continuation.*) Suppose $X \subseteq \prod X_i$ is finite. Prove that $\prec$ on X is irreflexive and condition E holds (for $m = 1, 2, \ldots$) if and only if there are real-valued functions $u_1, \ldots, u_n$ on $X_1, \ldots, X_n$ respectively such that, for all $x, y \in X$,

$$x \prec y \Leftrightarrow \sum_{i=1}^{n} u_i(x_i) + 1 < \sum_{i=1}^{n} u_i(y_i).$$

18. Scott (1964): Proof of Theorem 2.9. With $X/\approx$ finite select one element from each $\approx$ class and call the resulting set Y. Henceforth, work with Y. Each $x \prec y$ statement translates into $u(y) - u(x) - 1 > 0$ by (2.20), and each $x \sim y$ statement translates into $u(x) + 1 - u(y) > 0$ and $u(y) + 1 - u(x) > 0$. [≥ 0 might be used in the latter two, but >0 will work also.] Let N equal the size of Y plus 1, with $c = (u(x), \ldots, u(t), 1)$ being N-dimensional.

a. Use Theorem 4.2 to show that if there is no c solution to the stated inequalities then there are sequences $x_1, \ldots, x_T, z_1, \ldots, z_T$, and $y_1, \ldots, y_T, w_1, \ldots, w_T$ such that each is a permutation of the other and $x_k \prec y_k$, $z_k \sim w_k$ for $k = 1, \ldots, T$.

b. Show that $T = 1$ is impossible under the semiorder axioms.

c. Consider $T > 1$. Form a cycle through the two sequences by starting with some x_k. y_k is the second element in the cycle. Find y_k in the first sequence. Then the third element in the cycle is the element in the second sequence under y_k in the first. Continue this until you reach x_k in the second sequence. Show that if any such cycle stays wholly in the x_k, y_k pairs then transitivity of $\prec$ is violated.

d. Hence, with $T > 1$, a cycle beginning with x_k must pass through z_k, w_k pairs. Suppose some $y_j = x_k$. Then use $p11$ of Section 2.4 to show that you can reduce, by deletion and rearrangement, the two T sequences to $T-1$ sequences (one of which is a permutation of the other, with $T-1$ $\prec$ and $T-1$ $\sim$ statements

between the two). Suppose no $y_j = x_k$. Use $p10$ of Section 2.4 to show that the two T sequences can be reduced to corresponding $T - 1$ sequences.

e. Conclude the proof of Theorem 2.9.

19. Verify that $<^L$ on a set of n-dimensional real vectors is a strict order. (See Definition 2.1*b*.)

20. When Theorem 4.1C^* holds with X finite, discuss the nature of transformations on the u_i under which C^* will remain valid. Do the same for lexicographic utilities when (4.12) holds.

Chapter 5

ADDITIVE UTILITIES WITH INFINITE SETS

This chapter presents two well-structured theories for additive utilities on infinite sets. The earlier theory, due to Debreu (1960), is presented in Section 5.4. It is based on topological notions that are defined in Section 5.3. The other theory, due to Luce and Tukey (1964) and Luce (1966), is given in Section 5.2. As Krantz (1964) has noted, proofs in the latter theory can be based on the theory of ordered groups. Section 5.1 presents some of this theory.

Throughout this chapter, X is a complete Cartesian product, $X = \prod_{i=1}^{n} X_i$, and $<$ is assumed to be a weak order. Partly as a result of these assumptions along with a rather "tight" structure for $<$ on X, we shall not require all of condition C of Theorem 4.1. When $n = 2$, C_3 (condition C with $m = 3$) will suffice, and when $n \geq 3$, C_2 as in Theorem 4.3 will do. The assumptions of the theories imply the existence of additive utilities that are *unique up to similar positive linear transformations*. By this we mean that if real-valued functions $u_1, \ldots, u_n$ on $X_1, \ldots, X_n$ satisfy $x < y \Leftrightarrow \sum_i u_i(x_i) < \sum_i u_i(y_i)$, for all $x, y \in X$, then real-valued functions $v_1, \ldots, v_n$ on $X_1, \ldots, X_n$ satisfy $x < y \Leftrightarrow \sum_i v_i(x_i) < \sum_i v_i(y_i)$, for all $x, y \in X$, if and only if there are numbers $a, b_1, \ldots, b_n$ with $a > 0$ such that

$$v_i(x_i) = au_i(x_i) + b_i \qquad \text{for all } x_i \in X_i;\ i = 1, \ldots, n. \tag{5.1}$$

5.1 STRICTLY ORDERED GROUPS

A *group* is a set Y and a function that maps each $(x, y) \in Y \times Y$ into an element $x + y$ in Y such that for every $x, y, z \in Y$ and some fixed element $e \in Y$,

$G1$. $(x + y) + z = x + (y + z)$ (associativity)

$G2$. $x + e = e + x = x$ (identity)

$G3$. there is $-x \in Y$ such that $x + (-x) = -x + x = e$. (additive inverse)

e is the group *identity* and $-x$ is the *inverse* of x. A group $(Y, +)$ is *commutative* if the following holds throughout Y:

G4. $x + y = y + x$.

(Re, $+$) with $+$ natural addition and $e = 0$ is a commutative group. So is $(\{0, 1\}, +)$ with $-0 = 0, -1 = 1, 0 + 0 = 1 + 1 = 0, 0 + 1 = 1 + 0 = 1$ and $e = 0$.

When m is a positive integer, $mx = x + x + \cdots + x$ (m times). When m is a negative integer, $mx = -x - x - \cdots - x$ ($-m$ times). $0x = e$. If $(Y, +)$ is a group and m and n are integers, it is not hard to show that $mx + nx = (m + n)x$.

Definition 5.1. A *strictly ordered group* $(Y, +, <)$ is a group $(Y, +)$ and a strict order $<$ on Y such that, for all $x, y, z \in Y$,

$$x < y \Rightarrow x + z < y + z \quad \text{and} \quad z + x < z + y. \tag{5.2}$$

A strictly ordered group is *Archimedean* if and only if for all $x, y \in Y$, $(e < x, e < y) \Rightarrow y < mx$ for some positive integer m.

Let $Y = \{(j, k) : j$ and k are integers$\}$, let $+$ be natural addition, and let $< \, = \, <^L$, so that $(j, k) < (j', k') \Leftrightarrow j < j'$ or $(j = j', k < k')$. Then $(Y, +, <)$ is a strictly ordered group, but it is *not* Archimedean since $(0, 0) < (1, 0)$ and $(0, 0) < (0, 1)$ and $m(0, 1) = (0, m) < (1, 0)$ for every positive integer m. However, additive utilities exist for this case (Exercises 1*c*, 2). On the other hand if $Y = \{(r, s) : r$ and s are rational numbers$\}$ then again $(Y, +, <^L)$ is a non-Archimedean strictly ordered group but additive utilities do not exist for this case (Exercise 1*b*).

The following theorem, due to Hölder (1901), is used in the next section. The proof given is similar to Fuchs' proof (1963, pp. 45–46).

THEOREM 5.1. *Suppose that $(Y, +, <)$ is a strictly ordered group. Then $(Y, +, <)$ is Archimedean if and only if there is a real-valued function f on Y such that, for all $x, y \in Y$,*

$$x < y \Leftrightarrow f(x) < f(y) \tag{5.3}$$

$$f(x + y) = f(x) + f(y). \tag{5.4}$$

Moreover, if (5.3) *and* (5.4) *hold and if a real-valued function g on Y also preserves order* [*as in* (5.3)] *and is additive* [*as in* (5.4)] *then there is a real number $c > 0$ such that*

$$g(x) = cf(x) \qquad \text{for all } x \in Y, \tag{5.5}$$

and c is unique if $e < x$ for some $x \in Y$.

Proof. The fact that (5.3) and (5.4) imply the Archimedean property follows from $f(e) = 0$, using $G2$. To show the converse we assume that the Archimedean property holds and consider two exhaustive cases.

Case 1: set Y has a smallest "positive" element x so that $e \prec x$ and $e \prec y \prec x$ for no $y \in Y$. By the Archimedean property and $e \prec x \prec 2x \prec \cdots$, $0 \prec y$ implies that $mx \preceqq y \prec (m + 1)x$ for some positive integer m, where $z \preceqq y \Leftrightarrow z \prec y$ or $z = y$. Therefore $e \preceqq y - mx \prec x$ by (5.2), $G3$, $G2$, and $(m + 1)x = mx + x$. But then, by hypothesis for this case, $e = y - mx$, and thus $y = mx$ by (5.2), $G3$, and $G2$. Likewise, if $y \prec e$ then $y = mx$ for some negative integer m. Let $f(y) = m$ when $y = mx$. If $(y = m_1x, z = m_2x)$ then $y \prec z \Leftrightarrow m_1 < m_2$, so that (5.3) holds, and $f(y + z) = f(m_1x + m_2x) = f((m_1 + m_2)x) = m_1 + m_2$, verifying (5.4).

Case 2: if $e \prec x$ then $e \prec y \prec x$ for some $y \in Y$. For this case we first establish $G4$ (commutativity). Suppose $e \prec y \prec x$. Then either $2y \preceqq x$ or $x \prec 2y$. In the latter case $x - y \prec y$ by (5.2) and $2y - y = y$, so that $(x - y) + (x - y) \prec (x - y) + y$ by (5.2) and hence $2(x - y) \prec x$ by $G1$, $G3$, and $G2$. Moreover, $e \prec x - y$ by (5.2) and $G3$, and $y - x \prec x$ since $y - x \prec e$ and $e \prec x$. It follows that if $e \prec x$ then there is a $z \in Y$ such that $e \prec z \prec x$ and $2z \preceqq x$. Now suppose that Y is not commutative: for definiteness assume that $e \prec a$, $e \prec b$, and $a + b \neq b + a$ with $b + a \prec a + b$. Then let $x = (a + b) - (b + a)$ so that $e \prec x$ by (5.2) and $G3$, and let z be such that $e \succ z \prec x$ and $2z \preceqq x$ as just established. By the Archimedean property $(mz \preceqq a \prec (m + 1)z, nz \preceqq b \prec (n + 1)z)$ for non-negative integers m and n. Hence $a + b \prec (m + 1)z + b \prec (m + 1)z + (n + 1)z = (m + n + 2)z$ and $(n + m)z = nz + mz \preceqq b + mz \preceqq b + a$, or $-(b + a) \preceqq -(n + m)z$, so that $x = (a + b) - (b + a) \prec (m + n + 2)z - (n + m)z = 2z$, or $x \prec 2z$ thus contradicting $2z \preceqq x$. Hence $G4$ holds.

For Case 2 f is defined as follows, assuming $e \prec x$ for some $x \in Y$ to avoid the trivial situation. Fix a with $e \prec a$ and set $f(a) = 1$. For $x \in Y$ let

$$L_x = \{m/n : ma \preceqq nx,\ m \& n \text{ integers with } n > 0\}$$

$$U_x = \{m/n : nx \prec ma,\ m \& n \text{ integers with } n > 0\}.$$

$\{L_x, U_x\}$ is a partition of the rational numbers with $m/n < r/s$ whenever $m/n \in L_x$ and $r/s \in U_x$, as is easily seen. (For example, if $e \prec x$ then $ma \preceqq nx \Rightarrow sma \preceqq snx$ and $sx \prec ra \Rightarrow nsx \prec nra$ so that $sma \prec nra$, or $sm < nr$ (since $nr > 0$), or $m/n < r/s$.) It follows that there is a unique real number $f(x)$ such that

$$f(x) = \sup L_x = \inf U_x.$$

To prove that $f(x + y) = f(x) + f(y)$ suppose first that $m/n \in L_x$ and $r/s \in L_y$. Then $ma \preceqq nx$ and $ra \preceqq sy$. Hence $sma \preceqq snx$ and $nra \preceqq nsy$ so that $(ms + nr)a \preceqq ns(x + y)$, where $nsx + nsy = ns(x + y)$ on using $G4$

repeatedly to get $nsx + nsy = x + y + x + y + \cdots + x + y$. Therefore $(ms + nr)/ns = (m/n) + (r/s)$ is in L_{x+y}. Similarly, if $m/n \in U_x$ and $r/s \in U_y$ then $m/n + r/s$ is in U_{x+y}. It follows that

$$\sup L_x + \sup L_y \leq \sup L_{x+y} = f(x + y) = \inf U_{x+y} \leq \inf U_x + \inf U_y$$

and hence that $f(x + y) = \sup L_x + \sup L_y = f(x) + f(y)$. This proves (5.4).

To establish (5.3) suppose $e \prec x$. Then $a \prec mx$ for some positive m and hence $1/m \in L_x$ so that $f(x) > 0$. Similarly if $x \prec e$ then $f(-x) > 0$, and $f(e) = 0$ by $G2$ and (5.4). Hence $e \prec x \Leftrightarrow 0 < f(x)$, which is easily seen to imply (5.3).

The final part of the theorem, namely (5.5), is proved as follows. If $Y = \{e\}$ then $f(e) = g(e) = 0$ and every c satisfies (5.5). Next, suppose that $e \prec x$ for some $x \in Y$. If Case 1 above holds then, with $e \prec x$ and $e \prec y \prec x$ for no $y \in Y$, $f(z) = mf(x)$ and $g(z) = mg(x)$ when $z = mx$ so that $g(z) = [g(x)/f(x)]f(z)$ for all $z \in Y$. On the other hand suppose Case 2 holds with $e \prec a$. Then, by (5.3) and (5.4), $mf(a) \leq nf(x)$, $mg(a) \leq ng(x)$, $sf(x) \leq rf(a)$ and $sg(x) \leq rg(a)$ for all $m/n \in L_x$ and $r/s \in U_x$, from which it follows that $f(x)/f(a) = g(x)/g(a)$, or $g(x) = [g(a)/f(a)]f(x)$ for all $x \in Y$. ◆

5.2 ALGEBRAIC THEORY FOR n FACTORS

The additive-measurement theory developed by Luce and Tukey (1964) and Luce (1966) is based on the idea that a difference in two levels of one factor can be offset by a compensating difference in the levels of any other factor. For example, given $x_1^0 \in X_1$ and $x_2^0, x_2^1 \in X_2$, the compensation or "solvability" assumption says that $(x_1^0, x_2^1) \sim (x_1^1, x_2^0)$ for some $x_1^1 \in X_1$. If $X = X_1 \times X_2$ then $(x_1^1, x_2^1) \in X$ and again by solvability $(x_1^1, x_2^1) \sim (x_1^2, x_2^0)$ for some $x_1^2 \in X_1$. Under the cited conditions this gives rise to the picture in Figure 5.1 where the broken curves represent indifference sets. Suppose

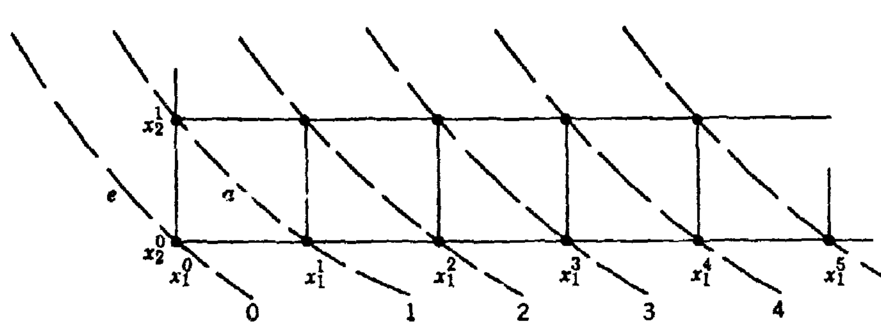

Figure 5.1 $X = X_1 \times X_2$.

additive utilities exist for this two-factor case and that, for points on e, such as (x_1^0, x_2^0), $u_1(x_1^0) + u_2(x_2^0) = 0$, and for points on a, such as (x_1^1, x_2^0), $u_1(x_1^1) + u_2(x_2^0) = 1$, with $(x_1^0, x_2^0) < (x_1^1, x_2^0)$. Then, as is easily verified, the value of $u_1 + u_2$ for the first curve to the right of a must be 2, for the next curve $u_1 + u_2 = 3$, and so forth. Thus, if $x < y \Leftrightarrow u_1(x_1) + u_2(x_2) < u_1(y_1) + u_2(y_2)$, then, for any $y \in X$ there must be a positive integer k such that $y < (x_1^k, x_2^0)$. Hence, under unrestricted solvability, we have a necessary Archimedean axiom for the two-factor case. It is $P3$ in Theorem 5.2.

Two Factors

In the following Theorem $P1$ (C_3 of Theorem 4.1) and $P3$ are necessary conditions for weak-order additivity when $X = X_1 \times X_2$, but unrestricted solvability ($P2$) is not. Except in the trivial case when $X/\sim\, = \{X\}$, $P2$ requires both u_1 and u_2 to be unbounded above and below. Luce (1966) shows how to weaken $P2$ to avoid the unboundedness implication: see also Krantz (1967, pp. 25–27).

THEOREM 5.2. *Suppose $X = X_1 \times X_2$ and the following three conditions hold throughout X:*

$P1$. $[(x^1, x^2, x^3)\, E_3\, (y^1, y^2, y^3), x^j < y^j$ *or* $x^j \sim y^j$ *for* $j < 3] \Rightarrow$ *not* $x^3 < y^3$.

$P2$. $(x_1, y_1 \in X_1; x_2 \in X_2) \Rightarrow (x_1, x_2) \sim (y_1, y_2)$ *for some* $y_2 \in X_2$, *and* $(x_1 \in X_1; x_2, y_2 \in X_2) \Rightarrow (x_1, x_2) \sim (y_1, y_2)$ *for some* $y_1 \in X_1$.

$P3$. $[(x_1^0, x_2^0) < (x_1^0, x_2^1),\ (x_1^{k-1}, x_2^1) \sim (x_1^k, x_2^0)$ *for* $k = 1, 2, \ldots; y \in X] \Rightarrow y < (x_1^k, x_2^0)$ *for some* $k \in \{1, 2, \ldots\}$.

Then there are real-valued functions u_1 on X_1 and u_2 on X_2 such that

$$x < y \Leftrightarrow u_1(x_1) + u_2(x_2) < u_1(y_1) + u_2(y_2), \qquad \textit{for all } x, y \in X, \tag{5.6}$$

and u_1 and u_2 satisfying (5.6) *are unique up to similar positive linear transformations.*

Proof. $P1$ implies that $<$ is a weak order (asymmetric, negatively transitive) so that $\sim$ is an equivalence. Let $X/\sim$ be the set of equivalence classes of X under $\sim$ and fix $(x_1^0, x_2^0) \in X$. By $P2$, each element in $X/\sim$ contains elements in X of the form (x_1, x_2^0), (x_1^0, x_2). Define $+$ on $X/\sim$ as follows:

$$\text{with } a, b \in X/\sim,\ a + b \text{ is the element in } X/\sim \text{ that contains } (x_1, x_2) \text{ when } (x_1, x_2^0) \in a \text{ and } (x_1^0, x_2) \in b. \tag{5.7}$$

With $a <' b \Leftrightarrow x < y$ for some $x \in a$ and $y \in b$, we first verify that $(X/\sim, +, <')$ is a strictly ordered commutative group. We then show that it is Archimedean and use Theorem 5.1.

1. $+$ is well defined. By $P1$, (x_1, x_2^0), $(y_1, x_2^0) \in a$ and (x_1^0, x_2), $(x_1^0, y_2) \in b$ imply that $(x_1, x_2) \sim (y_1, y_2)$.

2. Commutativity, $G4$. By $P1$, (x_1, x_2^0), $(x_1^0, y_2) \in a$ and (x_1^0, x_2), $(y_1, x_2^0) \in b \Rightarrow (x_1, x_2) \sim (y_1, y_2)$, and hence by (5.7), $a + b = b + a$.

3. Associativity, $G1$. By $G4$, $(a + b) + c = a + (b + c) \Leftrightarrow c + (a + b) = a + (c + b)$. Let $(x_1, x_2^0) \in a$, $(x_1^0, x_2) \in b$, $(y_1, x_2^0) \in c$, $(x_1^0, y_2) \in a + b$, and $(x_1^0, z_2) \in c + b$. Hence $(x_1^0, y_2) \sim (x_1, x_2)$ and $(y_1, x_2) \sim (x_1^0, z_2)$. Hence, by $P1$, $(y_1, y_2) \sim (x_1, z_2)$, which yields $c + (a + b) = a + (c + b)$.

4. Identity, $G2$. Let e contain (x_1^0, x_2^0). By $G4$, $e + a = a + e$. With $(x_1, x_2^0) \in a$, (5.7) implies $(x_1, x_2^0) \in a + e$. Hence $a = a + e$.

5. Additive Inverse, $G3$. Define $-a$ as that element in $X/\sim$ that contains (x_1^0, x_2) when $(x_1, x_2^0) \in a$ and $(x_1, x_2) \in e$. Then, by (5.7), $-a + a = e$.

6. $<'$ on $X/\sim$ is a strict order by Theorem 2.1. Suppose $a <' b$. With $(x_1, x_2^0) \in a$, $(y_1, x_2^0) \in b$, and $(x_1^0, x_2) \in c$, let z_1, by $P2$, satisfy $(z_1, x_2) \sim (y_1, x_2^0)$. With $(x_1, x_2^0) < (y_1, x_2^0)$ also, Theorem 2.1 yields $(x_1, x_2^0) < (z_1, x_2)$, which along with $(z_1, x_2) \sim (y_1, x_2^0)$ under $P1$ yields not $(y_1, x_2) < (x_1, x_2)$ and in fact $(x_1, x_2) < (y_1, x_2)$ since $(x_1, x_2) \sim (y_1, x_2)$ gives a violation of $P1$. Hence $a + c < b + c$, so that (5.2) holds.

To prove that $(X/\sim, +, <')$ is Archimedean, suppose $(e < a, e < b)$. With $(x_1^0, x_2^1) \in a$, let the sequence in $P3$ be constructed as described in connection with Figure 5.1. Since $(x_1^0, x_2^1) \in a$ and $(x_1^1, x_2^0) \in a$, (5.7) says that $(x_1^1, x_2^1) \in 2a$. Then, since $(x_1^2, x_2^0) \sim (x_1^1, x_2^1)$, $(x_1^2, x_2^0) \in 2a$. Using (5.7) to continue this we see that $(x_1^k, x_2^0) \in ka$, $k = 1, 2, \ldots$. With $y \in b$, $P3$ says that $y < (x_1^k, x_2^0)$ for some k, which gives $b <' ka$ for some positive integer k. Hence $(X/\sim, +, <')$ is Archimedean.

Thus, by Theorem 5.1, there is a real-valued function f on $X/\sim$ such that $f(a + b) = f(a) + f(b)$ and $a <' b \Leftrightarrow f(a) < f(b)$. Defining $u_1(x_1) = f(a)$ when $(x_1, x_2^0) \in a$, and $u_2(x_2) = f(b)$ when $(x_1^0, x_2) \in b$, (5.6) follows easily.

Suppose v_1 on X_1 and v_2 on X_2 also satisfy (5.6). Defining g on $X/\sim$ by $g(a) = [v_1(x_1) - v_1(x_1^0)] + [v_2(x_2) - v_2(x_2^0)]$ when $(x_1, x_2) \in a$, it follows that, taking $(x_1, x_2^0) \in a$ and $(x_1^0, x_2) \in b$ so that $(x_1, x_2) \in a + b$, $g(a) + g(b) = [v_1(x_1) - v_1(x_1^0)] + 0 + 0 + [v_2(x_2) - v_2(x_2^0)] = g(a + b)$. Moreover, from this and (5.6), $g(a) < g(b) \Leftrightarrow a <' b$. Hence, by Theorem 5.1, $g = cf$ for some positive number c. It follows that, taking (x_1, x_2^0), $v_1(x_1) - v_1(x_1^0) = cu_1(x_1)$, or $v_1(x_1) = cu_1(x_1) + v_1(x_1^0)$ for all $x_1 \in X_1$. Similarly, $v_2(x_2) = cu_2(x_2) + v_2(x_2^0)$ for all $x_2 \in X_2$. ◆

Three or More Factors

We now consider a version of Luce's theory (1966) for more than two factors. As pointed out to me by David Krantz (correspondence), the independence condition C_3 can be replaced by C_2 in this case. This necessitates of course the explicit assumption that $<$ is a weak order (or negatively

transitive) since weak order does not follow from C_3, or $P1^*$ as we call it below.

THEOREM 5.3. *Suppose* $X = \prod_{i=1}^n X_i$, $n \geq 3$, $<$ *on* X *is a weak order, and the following hold throughout* X:

$P1^*$. $[(x, z)\, E_2\, (y, w),\ x < y$ *or* $x \sim y] \Rightarrow$ *not* $z < w$.

$P2^*$. $[i \in \{1, \ldots, n\},\ x \in X$ *and* $y_j \in X_j$ *for all* $j \neq i] \Rightarrow x \sim (y_1, \ldots, y_{i-1}, z_i, y_{i+1}, \ldots, y_n)$ *for some* $z_i \in X_i$.

$P3^*$. $[(x_1^0, x_2^0, \ldots, x_n^0) < (x_1^0, x_2^1, \ldots, x_n^1),\ (x_1^{k-1}, x_2^1, \ldots, x_n^1) \sim (x_1^k, x_2^0, \ldots, x_n^0)$ *for* $k = 1, 2, \ldots;\ y \in X] \Rightarrow y < (x_1^k, x_2^0, \ldots, x_n^0)$ *for some* $k \in \{1, 2, \ldots\}$.

Then there are real-valued functions $u_1, \ldots, u_n$ *on* $X_1, \ldots, X_n$ *respectively such that*

$$x < y \Leftrightarrow \sum_{i=1}^n u_i(x_i) < \sum_{i=1}^n u_i(y_i), \qquad \text{for all } x, y \in X, \tag{5.8}$$

and $u_1, \ldots, u_n$ *satisfying* (5.8) *are unique up to similar positive linear transformations.*

Proof. Our major task will be to show that C_3 or $P1$ for $n \geq 3$ follows from the stated hypotheses. We delay this until later, assuming for the moment that C_3 or $P1$ holds. Fix $(x_1^0, \ldots, x_n^0)$. By $P2^*$ any $a \in X/\sim$ contains elements of the form (x_i for $i \in I$, x_i^0 for $i \notin I$) for any nonempty proper subset $I \subset \{1, \ldots, n\}$. Define $+$ on $X/\sim$ as follows:

> $a + b$ is the element in $X/\sim$ that contains $(x_1, \ldots, x_n)$ when, for any nonempty $I \subset \{1, \ldots, n\}$, ($x_i$ for $i \in I$, x_i^0 for $i \notin I) \in a$ and (x_i^0 for $i \in I$, x_i for $i \notin I) \in b$.

To show that $+$ is well defined suppose (x_i for $i \in I$, x_i^0 for $i \notin I) \in a$, (x_i^0 for $i \in I$, x_i for $i \notin I) \in b$, (y_i for $i \in I^*$, x_i^0 for $i \notin I^*) \in a$, and (x_i^0 for $i \in I^*$, y_i for $i \notin I^*) \in b$, when I and I^* are any two nonempty proper subsets of $\{1, \ldots, n\}$. We need to prove that $(x_1, \ldots, x_n) \sim (y_1, \ldots, y_n)$, and this is easily seen to follow from $P1$.

By analogy with the preceding proof (let X_2 there represent $X_2 \times \cdots \times X_n$ here) it follows from $P1$ and the hypotheses of Theorem 5.3 that $(X/\sim, +, <')$ is an Archimedean simply ordered (commutative) group. With f as in (5.3) and (5.4) and $(x_1^0, \ldots, x_n^0) \in e$, define $u_i(x_i) = f(a)$ when $(x_1^0, \ldots, x_{i-1}^0, x_i, x_{i+1}^0, \ldots, x_n^0) \in a$, and define $u(x) = f(a)$ when $x \in a$. Then $u(x) < u(y) \Leftrightarrow f(a) < f(b)$ so that $x < y \Leftrightarrow u(x) < u(y)$. Moreover, with $(x)^\sim$ the element in $X/\sim$ that contains x, it follows from successive uses of our definition for $+$ that $(x_1, \ldots, x_n)^\sim = (x_1, x_2^0, \ldots, x_n^0)^\sim + (x_1^0, x_2, \ldots, x_n)^\sim = (x_1, x_2^0, \ldots, x_n^0)^\sim + [(x_1^0, x_2, x_3^0, \ldots, x_n^0)^\sim + (x_1^0, x_2^0, x_3, \ldots, x_n)^\sim] = \cdots = (x_1, x_2^0, \ldots, x_n^0)^\sim + (x_1^0, x_2, x_3^0, \ldots, x_n^0)^\sim + \cdots + (x_1^0, \ldots, x_{n-1}^0, x_n)^\sim$, from which we

obtain $u(x) = u_1(x_1) + u_2(x_2) + \cdots + u_n(x_n)$. The proof of uniqueness follows from Theorem 5.1 as in the preceding proof.

Proof of P1. Let $X = \prod_{i=1}^n X_i$, $n \geq 3$, and assume that $P1^*$ and $P2^*$ hold and that $<$ is a weak order. To verify $P1$ we begin with the following general form: show that $(x_1, x_2, \ldots, x_6) \lesssim (y_1, y_2, y_3, y_4, x_5, x_6)$ when

$$(x_1, x_2, z_3, z_4, z_5, z_6) \lesssim (y_1, z_2, y_3, z_4, t_5, z_6) \tag{5.9}$$

$$(z_1, z_2, x_3, x_4, t_5, t_6) \lesssim (z_1, y_2, z_3, y_4, z_5, t_6). \tag{5.10}$$

This includes all possible placements of x_i, y_i, etc. in the two given statements. It should be understood that some dimensions may be collected into a single i in (5.9) and (5.10) and that one or more of the i patterns in (5.9) or (5.10) may be absent in a specific case.

Suppose first that the first dimension ($i = 1$) in (5.9) and (5.10) is actually present. Using $P2^*$ let s_1 satisfy (omitting parentheses and commas) $s_1z_2z_3z_4t_5z_6 \sim x_1x_2z_3z_4z_5z_6$. Then, by $P1^*$, $x_1x_2x_3x_4z_5z_6 \sim s_1z_2x_3x_4t_5z_6$. Also, since $s_1z_2z_3z_4t_5z_6 \lesssim y_1z_2y_3z_4t_5z_6$ by (5.9), $P1^*$ implies $s_1y_2z_3y_4z_5z_6 \lesssim y_1y_2y_3y_4z_5z_6$. Also, by (5.10) and $P1^*$, $s_1z_2x_3x_4t_5z_6 \lesssim s_1y_2z_3y_4z_5z_6$. Hence, by transitivity, $x_1x_2x_3x_4z_5z_6 \lesssim y_1y_2y_3y_4z_5z_6$, so that by $P1^*$ $x_1x_2x_3x_4x_5x_6 \lesssim y_1y_2y_3y_4x_5x_6$.

The key to this proof was that the same element (z_1) appeared in the first position on each side of (5.10). A similar proof holds if either the fourth dimension is present (z_4 on both sides of (5.9)) or the sixth dimension (z_6) is present. Assume henceforth that none of these three dimensions is actually present. Renumbering subscripts, (5.9) and (5.10) then reduce to

$$(x_1, z_2, z_3) \lesssim (z_1, y_2, t_3) \tag{5.11}$$

$$(z_1, x_2, t_3) \lesssim (y_1, z_2, z_3). \tag{5.12}$$

We are to show that $x_1x_2x_3 \lesssim y_1y_2x_3$. Assuming that the third dimension is present let s_3 satisfy $x_1z_2z_3 \sim z_1z_2s_3$. Then, by (5.11) and $P1^*$, $y_1z_2s_3 \lesssim y_1y_2t_3$. Also, $x_1z_2z_3 \sim z_1z_2s_3$ and (5.12) satisfy the condition in the preceding proof and can conform to E_3 so that, by $P1$ for this case $x_1x_2t_3 \lesssim y_1z_2s_3$. Then, by transitivity, $x_1x_2t_3 \lesssim y_1y_2t_3$ so that $x_1x_2x_3 \lesssim y_1y_2x_3$ by $P1^*$.

Finally, suppose that the third dimension in the preceding paragraph (fifth dimension originally) is not present. This leaves us with only two patterns. But $n \geq 3$. Therefore, we have a case like

$$(x_1, z_2, z_3) \lesssim (z_1, y_2, y_3) \tag{5.13}$$

$$(z_1, x_2, x_3) \lesssim (y_1, z_2, z_3) \tag{5.14}$$

from which we are to show that $x_1x_2x_3 \lesssim y_1y_2y_3$. Let s_3 satisfy $x_1z_2z_3 \sim z_1z_2s_3$. By (5.13) and $P1^*$, $y_1z_2s_3 \lesssim y_1y_2y_3$. Also, $x_1z_2z_3 \sim z_1z_2s_3$ and (5.14) satisfy

the previous pattern for which $P1$ holds (z_2 on both sides of the $\sim$ statement), so that by $P1$ for this case, $x_1x_2x_3 \precsim y_1z_2s_3$. Then, by transitivity, $x_1x_2x_3 \precsim y_1y_2y_3$. ◆

5.3 TOPOLOGICAL PRELIMINARIES

To obtain a sound understanding of Debreu's (1960) additivity theory a review of some theory of topology is in order. Familiarity with Section 3.4 is assumed.

A topological space $(X, \mathcal{T})$ is *connected* if and only if X cannot be partitioned into two nonempty open sets (in $\mathcal{T}$). The *closure* $\bar{A}$ of $A \subseteq X$ is the set of all $y \in X$ for which every open set that contains y has a nonempty intersection with A:

$$\bar{A} = \{y : y \in X \quad \text{and} \quad (y \in B, B \in \mathcal{T}) \Rightarrow A \cap B \neq \varnothing\}. \tag{5.15}$$

$(X, \mathcal{T})$ is *separable* if and only if X includes a countable subset whose closure is X. $(\mathrm{Re}, \mathcal{U})$ is separable (as well as connected) since Re is the closure of the set of all rational numbers.

The following is Debreu's Proposition 4 (1964, p. 291).

LEMMA 5.1. *Suppose $\prec$ on X is a weak order, $(X, \mathcal{T})$ is a connected and separable topological space, and $\{x : x \in X, x \prec y\} \in \mathcal{T}$ and $\{x : x \in X, y \prec x\} \in \mathcal{T}$ for every $y \in X$. Then there is a real-valued function u on X that is continuous in the topology $\mathcal{T}$ and satisfies*

$$x \prec y \Leftrightarrow u(x) < u(y), \qquad \textit{for all } x, y \in X. \tag{5.16}$$

Proof. By separability, X includes a countable subset A with $\bar{A} = X$. If $x \prec z$ then $\{y : y \prec z\}$ and $\{y : x \prec y\}$ are nonempty intersecting (by connectedness) open sets with intersection $\{y : x \prec y \prec z\} \in \mathcal{T}$. Then, by $\bar{A} = X$ and (5.15), $\{y : x \prec y \prec z\} \cap A \neq \varnothing$. Hence A is $\prec$-order dense in X. Theorems 3.1 and 3.5 complete the proof. ◆

Lemma 5.1 is used in the next section. Lemma 5.2, based on the following definition, is used later in this section. Given a topological space $(X, \mathcal{T})$, $Y \subseteq X$ is *connected* if and only if $(Y \cap A \neq \varnothing, Y \cap B \neq \varnothing, Y \subseteq A \cup B, Y \cap A \cap B = \varnothing)$ is false for every $A, B \in \mathcal{T}$.

LEMMA 5.2. *If $A \subseteq X$ is connected for each $A \in \mathcal{A}$ and if $A \cap A^* \neq \varnothing$ when $A, A^* \in \mathcal{A}$ then $\bigcup_{\mathcal{A}} A$ is connected.*

Proof. Suppose $Y = \bigcup_{\mathcal{A}} A$ is not connected. Then $(Y \cap B \neq \varnothing, Y \cap C \neq \varnothing, Y \subseteq B \cup C, Y \cap B \cap C = \varnothing)$ for some $B, C \in \mathcal{T}$. Let

$A, A^* \in \mathcal{A}$ satisfy $(A \cap B \neq \varnothing, A^* \cap C \neq \varnothing)$. Using B and C it follows that $A \cup A^*$ is not connected and that, since each of A and A^* is connected by hypothesis, it must be true that $(A \cap C = \varnothing, A^* \cap B = \varnothing)$. But then $(A \subseteq B, A^* \subseteq C)$, since $A, A^* \subseteq Y \subseteq B \cup C$ and hence $\varnothing = A \cap C \cap A^* \cap B = (A \cap B) \cap (A^* \cap C) = A \cap A^*$, which contradicts our second hypothesis. ◆

Product Topologies

If $X = \prod_{i=1}^n X_i$ and $(X_i, \mathcal{T}_i)$ is a topological space for each i let

$$\prod \mathcal{T}_i = \left\{A : A \subseteq X \text{ and if } (x_1, \ldots, x_n) \in A \text{ then there are } A_i \in \mathcal{T}_i \text{ for which } x_i \in A_i \, (i = 1, \ldots, n) \text{ and } \prod_{i=1}^n A_i \subseteq A\right\}. \tag{5.17}$$

$\prod \mathcal{T}_i$ is the *product topology* for $X = \prod X_i$. The *product space* $\prod (X_i, \mathcal{T}_i)$ is the topological space $(\prod X_i, \prod \mathcal{T}_i)$.

To verify that $\prod \mathcal{T}_i$ is indeed a topology (Definition 3.2), we note first that $\varnothing \in \prod \mathcal{T}_i$ and $X \in \prod \mathcal{T}_i$. Let B be a union of sets in $\prod \mathcal{T}_i$ with $B \neq \varnothing$ and $x \in B$. Then $x \in A$ for some A in the union and in $\prod \mathcal{T}_i$ from which it follows that $B \in \prod \mathcal{T}_i$. Finally, suppose $A^1, \ldots, A^m \in \prod \mathcal{T}_i$ and $\bigcap A^j \neq \varnothing$. With $x \in \bigcap A^j$, there is for each i and j an $A_i^j \in \mathcal{T}_i$ such that $x_i \in A_i^j$ and $\prod_i A_i^j \subseteq A^j$. Since $x_i \in \bigcap_j A_i^j \in \mathcal{T}_i$ for each i and $\prod_i (\bigcap_j A_i^j) = \bigcap_j (\prod_i A_i^j) \subseteq \bigcap_j A^j$, it follows from (5.17) that $\bigcap A^j \in \prod \mathcal{T}_i$.

Exercise 13 gives an equivalent definition of a product topology.

LEMMA 5.3. *If $(X_i, \mathcal{T}_i)$ is a connected (separable) topological space for each $i \in \{1, \ldots, n\}$ then $(\prod X_i, \prod \mathcal{T}_i)$ is connected (separable).*

Proof. Separability. Let $(X_i, \mathcal{T}_i)$ be separable for each i, with $A_i \subseteq X_i$ countable and $\bar{A}_i = X_i$. Given $x \in X = \prod X_i$ suppose $x \in B \in \prod \mathcal{T}_i$. Then by (5.17) there are $B_i \in \mathcal{T}_i$ such that $x_i \in B_i$ and $\prod B_i \subseteq B$. Since $(X_i, \mathcal{T}_i)$ is separable, $B_i \cap A_i \neq \varnothing$. Therefore $(\prod B_i) \cap (\prod A_i) \neq \varnothing$ so that $B \cap (\prod A_i) \neq \varnothing$. It follows that x is in the closure of $\prod A_i$.

Connectedness. Using the definition of a connected subset it is not hard to show that $\{x_1\} \times \cdots \times \{x_{i-1}\} \times X_i \times \{x_{i+1}\} \times \cdots \times \{x_n\}$ is a connected subset of $X = \prod X_i$ when $(X_i, \mathcal{T}_i)$ is connected. Let each $(X_i, \mathcal{T}_i)$ be connected so that $X_1 \times \{x_2\} \times \cdots \times \{x_n\} \cup \{y_1\} \times X_2 \times \{x_3\} \times \cdots \times \{x_n\}$ is connected by Lemma 5.2 since $(y_1, x_2, \ldots, x_n)$ is in both parts of the union. Since $(x_1, x_2, \ldots, x_n)$ is in every such union as y_1 varies over X_1 it follows by Lemma 5.2 that $X_1 \times X_2 \times \{x_3\} \times \cdots \times \{x_n\}$ is connected. Hence $X_1 \times X_2 \times \{x_3\} \times \cdots \times \{x_n\} \cup \{y_1\} \times \{y_2\} \times X_3 \times \{x_4\} \times \cdots \times \{x_n\}$ is connected, so that $X_1 \times X_2 \times X_3 \times \{x_4\} \times \cdots \times \{x_n\}$ is connected. By induction, $X_1 \times X_2 \times \cdots \times X_n$ is connected. ◆

Continuity

The appropriate generalization of continuity over that given in Section 3.4 is included in the following definition.

Definition 5.2. Let $(X, \mathcal{R})$, $(Y, \mathcal{S})$, and $(Z, \mathcal{T})$ be topological spaces. If f is a function on X into Y then *f is $\mathcal{R} - \mathcal{S}$ continuous* if and only if $S \in \mathcal{S} \Rightarrow \{x : x \in X, f(x) \in S\} \in \mathcal{R}$. If g is a function on $X \times Y$ into Z then

1. *g is continuous in X* if and only if $(y \in Y, T \in \mathcal{T}) \Rightarrow \{x : x \in X, g(x, y) \in T\} \in \mathcal{R}$;
2. *g is continuous in Y* if and only if $(x \in X, T \in \mathcal{T}) \Rightarrow \{y : y \in Y, g(x, y) \in T\} \in \mathcal{S}$.

The following lemma is used in the next section.

LEMMA 5.4. *If $(X, \mathcal{R})$, $(Y, \mathcal{S})$, and $(Z, \mathcal{T})$ are topological spaces and f on $X \times Y$ into Z is $\mathcal{R} \times \mathcal{S} - \mathcal{T}$ continuous then f is continuous in X and in Y.*

Proof. Let f be $\mathcal{R} \times \mathcal{S} - \mathcal{T}$ continuous, and let $b \in Y$, $T \in \mathcal{T}$. We shall show that $\{x : x \in X, f(x, b) \in T\} \in \mathcal{R}$. For all $x \in X$ let $g(x) = x$, $h(x) = b$ and $k(x) = (g(x), h(x))$. As is easily verified, g on X into X is $\mathcal{R} - \mathcal{R}$ continuous and h on X into Y is $\mathcal{R} - \mathcal{S}$ continuous. To show that k on X into $X \times Y$ is $\mathcal{R} - \mathcal{R} \times \mathcal{S}$ continuous let $A \neq \varnothing$, $A \in \mathcal{R} \times \mathcal{S}$. By Exercise 13, A has the form

$$A = \bigcup_{w \in W} B(w) \times C(w)$$

with $B(w) \in \mathcal{R}$ and $C(w) \in \mathcal{S}$ for all $w \in W$. Letting a super -1 denote the inverse $[k^{-1}(A) = \{x : k(x) \in A\}]$, it follows that

$$\begin{aligned}
k^{-1}(A) &= k^{-1}(\bigcup [B(w) \times C(w)]) \\
&= \bigcup k^{-1}[B(w) \times C(w)] \\
&= \bigcup k^{-1}([B(w) \times Y] \cap [X \times C(w)]) \\
&= \bigcup (k^{-1}[B(w) \times Y] \cap k^{-1}[X \times C(w)]) \\
&= \bigcup (g^{-1}[B(w)] \cap h^{-1}[C(w)]) \in \mathcal{R}
\end{aligned}$$

since $g^{-1}[B(w)] \in \mathcal{R}$ and $h^{-1}[C(w)] \in \mathcal{R}$ for every w.

Let $r(x) = f(k(x)) = f(x, b)$. Let $T \in \mathcal{T}$. Since $f^{-1}(T) \in \mathcal{R} \times \mathcal{S}$, and $k^{-1}(f^{-1}(T)) \in \mathcal{R}$ by the preceding demonstration, $r^{-1}(T) = k^{-1}(f^{-1}(T))$ is in $\mathcal{R}$. That is, $\{x : x \in X, f(x, b) \in T\} \in \mathcal{R}$, as desired. ◆

Suppose $(X_i, \mathcal{T}_i)$ is a topological space for each i and u is a real-valued function on $X = \prod_{i=1}^{n} X_i$ that is $\prod \mathcal{T}_i - \mathcal{U}$ continuous (continuous in the

topology $\prod \mathcal{T}_i$). Then, as a corollary of Lemma 5.4, u is continuous in $\prod_{i\in I} \mathcal{T}_i$ for every nonempty $I \subseteq \{1, 2, \ldots, n\}$. That is, for each $U \in \mathcal{U}$ and fixed x_i^0 for $i \notin I$, $\{x_I : x \in \prod X_i, x_i = x_i^0 \text{ for all } i \notin I, u(x) \in U\} \in \prod_{i\in I} \mathcal{T}_i$.

5.4 TOPOLOGICAL THEORY FOR n FACTORS

Debreu's (1960) topologically oriented theory for additive utility with two factors is essentially as follows.

THEOREM 5.4. *Suppose $X = X_1 \times X_2$ and the following three conditions hold throughout X:*

Q1. $[(x^1, x^2, x^3)\, E_3\, (y^1, y^2, y^3), x^j < y^j$ *or* $x^j \sim y^j$ *for* $j < 3] \Rightarrow$ *not* $x^3 < y^3$.
Q2. $(X_i, \mathcal{T}_i)$ *is a connected and separable topological space for* $i = 1, 2$.
Q3. $\{x : x \in X, x < y\} \in \mathcal{T}_1 \times \mathcal{T}_2$ *and* $\{x : x \in X, y < x\} \in \mathcal{T}_1 \times \mathcal{T}_2$.

Then there are real-valued functions u_1 on X_1 and u_2 on X_2 that satisfy (5.6) *and, if $(x_1, x_2) < (x_1, y_2)$ and $(y_1, z_2) < (z_1, z_2)$ for some quartet of elements in X, then u_1 and u_2 satisfying* (5.6) *are continuous in $\mathcal{T}_1$ and $\mathcal{T}_2$ respectively and are unique up to similar positive linear transformations.*

The obvious difference between this and Theorem 5.2 is in the $Q2$ and $Q3$ conditions ($Q1 = P1$). Debreu ties the space together with topological conditions, whereas Luce and Tukey use solvability. The need for the quartet condition in Theorem 5.4 stems from the fact that under $Q1$, $Q2$, and $Q3$ it is possible to have, say, u_1 constant on X_1 and u_2 nonconstant on X_2, in which case additive utilities are not unique up to similar positive linear transformations and u_2 need not be continuous. With no loss in generality *we assume in what follows that $(x_1, x_2) < (x_1, y_2)$ and $(y_1, z_2) < (z_1, z_2)$ for some quartet of elements in X.*

The most obvious application of Theorem 5.4 arises when X_1 and X_2 are intervals in Re. In fact, Part I of the two-part proof of the theorem assumes that $X_1 \times X_2$ is a rectangular subset of Re^2. Part II then shows how the general case can be transformed into the plane. Because Part I, which involves ideas of Thomsen (1927) and Blaschke (1928) for what Debreu calls the Thomsen-Blaschke theorem, goes through many steps and is rather long, I shall not detail every step.

Proof, Part I. Throughout we assume that the hypotheses of the theorem hold, that X_1 and X_2 are nondegenerate intervals of real numbers, that $\mathcal{T}_1$ and $\mathcal{T}_2$ are the relative usual topologies on X_1 and X_2, and that

$$x < y \Rightarrow x < y \qquad (5.18)$$

as in Theorem 3.3, condition 2.

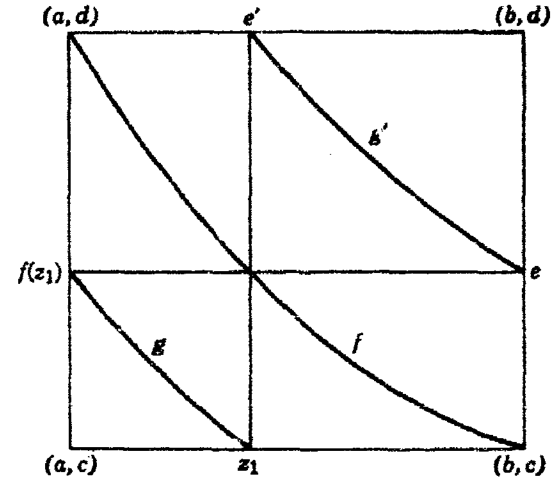

Figure 5.2

1. By Lemmas 5.3 and 5.1 there is a real-valued function v on X that is continuous in $\mathcal{T}_1 \times \mathcal{T}_2$ and satisfies $x \prec y \Leftrightarrow v(x) < v(y)$. Then, by Exercises 3.22 and 3.10,

$$(x \prec y, y \prec z, x < z) \Rightarrow y \sim \alpha x + (1 - \alpha)z \qquad \text{for a unique } \alpha \in (0, 1). \tag{5.19}$$

2. Let $[a, b] \times [c, d]$ be a rectangular subset of $X_1 \times X_2$ for which $(b, c) \sim (a, d)$. (Figure 5.2.) From (5.18) and (5.19) it follows that there is a real-valued function f on $[a, b]$ onto $[c, d]$ that is one-to-one with

$$(x_1, f(x_1)) \sim (b, c) \qquad \text{for every } x_1 \in [a, b].$$

Since f strictly decreases as x_1 increases, it and its inverse f^{-1} are continuous.

3. Our immediate goal is to show that additive utilities satisfying

$$u(x_1, x_2) = u_1(x_1) + u_2(x_2), \tag{5.20}$$

exist on $[a, b] \times [c, d]$ with u a monotonic transformation of v in step 1. First, set $u_1(a) = u_2(c) = 0$ and $u_1(b) = u_2(d) = 1$. Then $u(a, c) = 0$, $u(x_1, x_2) = 1$ for all $(x_1, x_2) \in f$, and $u(b, d) = 2$. As shown in Figure 5.2 there is a $z_1 \in (a, b)$ such that $(z_1, c) \sim (a, f(z_1))$. To prove this note first that since v is continuous it is continuous in X_1 and X_2 by Lemma 5.4. Then, by Exercise 3.16, $\{v(x_1, c): x_1 \in [a, b]\}$ is an interval in Re. Likewise, $\{v(a, f(x_1)): x_1 \in [a, b]\} = \{v(a, x_2): x_2 \in [c, d]\}$ is an interval. Since $v(x_1, c)$ increases in x_1 and $v(a, f(x_1))$ decreases in x_1, there is a unique $z_1 \in (a, b)$ for which $v(z_1, c) = v(a, f(z_1))$, so that $(z_1, c) \sim (a, f(z_1))$.

Let g be the continuous indifference curve through (z_1, c). To satisfy (5.20) we must have $u_1(z_1) = u_2(f(z_1)) = \frac{1}{2}$ and $u(x_1, x_2) = \frac{1}{2}$ for every $(x_1, x_2) \in g$.

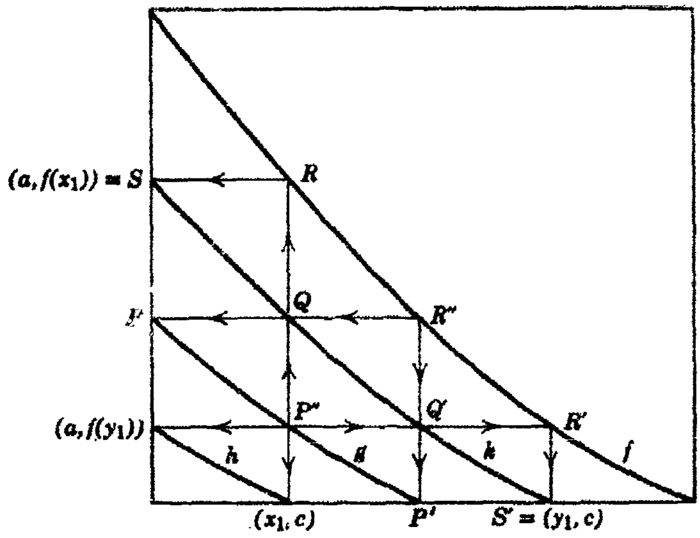

Figure 5.3

$Q1$ implies that $e \sim e'$ as shown in Figure 5.2, with $u(x_1, x_2) = \frac{3}{2}$ for every $(x_1, x_2) \in g'$.

For reasons like those given above there is a point $(x_1, f(y_1)) = P''$ in g for which $(x_1, c) \sim (a, f_1(y_1))$. As shown in Figure 5.3, the constructions from P'' define two new curves h and k. As is easily seen from $Q1$, $Q \sim Q'$, $Q \sim S'$, and $Q' \sim S$ so that Q, Q', S, and S' do indeed lie on the same indifference curve (k). For (5.20) we must set $u_1(x_1) = u_2(f(y_1)) = \frac{1}{4}$ and $u_1(y_1) = u_2(f(x_1)) = \frac{3}{4}$ with $u = \frac{1}{2}$ for h and $u = \frac{3}{2}$ for k. Similar constructions (from g' in Figure 5.2) hold above f on Figure 5.3.

4. The process of generating indifference curves in $[a, b] \times [c, d]$ is repeated *ad infinitum* and yields a continuous indifference curve for each value of u in $\{m/2^n : 1 \leq m \leq 2^n,\ n = 1, 2, \ldots\} \cup \{1 + m/2^n : 0 \leq m \leq 2^n - 1,\ n = 1, 2, \ldots\}$. If (x_1, x_2) and (y_1, y_2) are on these curves then $(x_1, x_2) < (y_1, y_2) \Leftrightarrow u(x_1, x_2) < u(y_1, y_2)$.

In addition, we have a set A of x_1 points in $[a, b]$ whose set of u_1 values is $\{m/2^n : 0 \leq m \leq 2^n,\ n = 1, 2, \ldots\}$ and a set B of x_2 points in $[c, d]$ whose set of u_2 values is $\{m/2^n : 0 \leq m \leq 2^n,\ n = 1, 2, \ldots\}$ with $u(x_1, x_2) = u_1(x_1) + u_2(x_2)$ whenever $(x_1, x_2) \in A \times B$.

5. $\bar{A} = [a, b]$ and $\bar{B} = [c, d]$. (We leave this closure proof to the reader.) It follows that

$$\sup \{u_1(y_1) : y_1 \leq x_1, y_1 \in A\} = \inf \{u_1(z_1) : x_1 \leq z_1, z_1 \in A\} \tag{5.21}$$

for each $x_1 \in [a, b]$. Extending u_1 on A to u_1 on $[a, b]$ by defining $u_1(x_1)$ as the common value in (5.21), it follows easily that $u_1(x_1) < u_1(y_1) \Leftrightarrow x_1 < y_1$ and that u_1 on $[a, b]$ is continuous. It is clear also that once $u_1(a)$ and $u_1(b)$ are

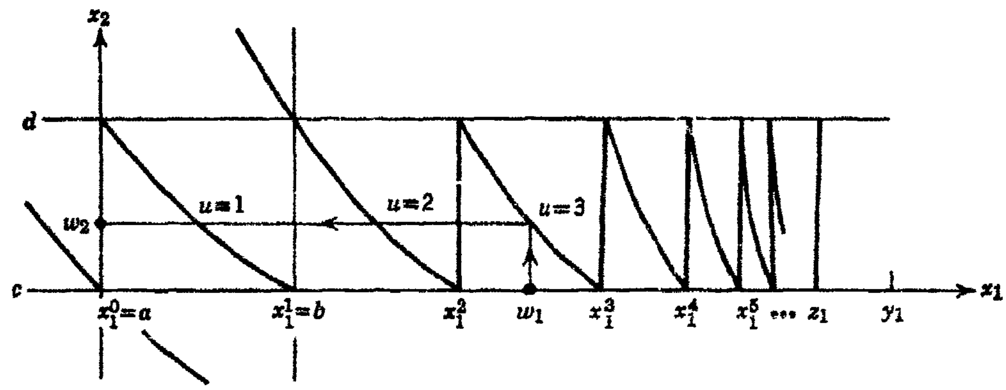

Figure 5.4

specified, the rest of u_1 on $[a, b]$ is uniquely determined and u_1 on $[a, b]$ must be continuous.

Similar remarks hold for u_2 on $[c, d]$, and, if u_1 and u_2 satisfy (5.20), they are unique up to similar positive linear transformations.

To verify additivity on $[a, b] \times [c, d]$ suppose first that $(x_1, x_2) \sim (y_1, y_2)$, both points being in $[a, b] \times [c, d]$. Since additivity holds on $A \times B$, we have from (5.21) and its companion for B that $u_1(x_1) + u_2(x_2) = u_1(y_1) + u_2(y_2)$. On the other hand suppose that $(x_1, x_2) < (y_1, y_2)$. Then, as is easily seen there must be a point $(z_1, z_2) \sim (x_1, x_2)$ for which $(z_1, z_2) < (y_1, y_2)$ so that $u_1(z_1) + u_2(z_2) < u_1(y_1) + u_2(y_2)$ and hence $u_1(x_1) + u_2(x_2) < u_1(y_1) + u_2(y_2)$.

6. We now show that u_1 and u_2 can be extended in one and only one way to all of X_1 and X_2 to satisfy additivity. Beginning with $[a, b] \times [c, d]$ we first extend the horizontal lines through (a, c) and (b, c) and through (a, d) and (b, d), and likewise for the two vertical lines. The indifference curves through (a, c), (b, c), and (b, d) are extended also. The procedure described in connection with Figure 5.1 is then used to generate additional indifference curves that must have u values of $2, 3, 4, \ldots$, and $-1, -2, \ldots$, this process continuing indefinitely or until the border(s) of X (if any) are reached. This provides us with a grid pattern on $X_1 \times X_2$ of rectangles similar to $[a, b] \times [c, d]$, except that some of these will be truncated if X is bounded. Using $Q1$ it is easy to verify that (except at the boundary) the lower right corner of any rectangle is indifferent to its upper left corner.

7. We need to show that these rectangles (including truncated ones at the boundaries, if any) actually cover $X_1 \times X_2$. For this it will suffice to show that every $x_1 > b$ lies beneath an indifference curve generated in the manner of Figure 5.1. To the contrary suppose, as in Figure 5.4, that $y_1 \in X_1$ does not satisfy this condition. Let $z_1 = \sup\{x_1^j : j = 0, 1, \ldots\}$ as shown on the figure. The continuity of v then implies that $v(z_1, c) = \sup\{v(x_1^j, c) : j = 0, 1, \ldots\}$ and $v(z_1, d) = \sup\{v(x_1^j, d) : j = 0, 1, \ldots\}$ so that $v(z_1, c) = v(z_1, d)$ and hence $(z_1, c) \sim (z_1, d)$, which contradicts (5.18).

8. For additivity it is clear that $u_1(x_1^j) = j$ for each such point on the x_1 axis of Figure 5.4. Suppose $w_1 \in X_1$: for definiteness we assume $w_1 > b$ as shown on Figure 5.4. By the construction shown for w_1, additivity requires $u_1(w_1) + u_2(w_2) = 3$. But $u_2(w_2)$ is already known since $w_2 \in [c, d]$. Hence $u_1(w_1)$ is uniquely determined. Similar remarks hold if $w_1 < a$, and, by symmetry, for points in X_2 not in $[c, d]$. Thus, given additive u_1 and u_2 on $[a, b]$ and $[c, d]$, u_1 and u_2 are uniquely determined on all of X_1 and X_2 when additivity is required, and they are continuous.

9. It remains to show that (5.6) holds throughout $X_1 \times X_2$. For this it will suffice to show that $(x_1, x_2) \sim (y_1, y_2) \Rightarrow u_1(x_1) + u_2(x_2) = u_1(y_1) + u_2(y_2)$ because then all points on the same indifference curve will have a common $u_1 + u_2$ value and, by construction, one such curve is to the left of another if and only if the former has a smaller $u_1 + u_2$ value.

We begin this with the rectangle in the grid of step 6 that is to the immediate right of $[a, b] \times [c, d]$. Suppose first that $x \sim y$ and these points are beneath the $u = 2$ curve as shown on Figure 5.5. By the constructions shown in the figure, and using Q1, $(P \sim P', z \sim x) \Rightarrow Q \sim Q'$ and $(P \sim P'', z \sim y) \Rightarrow R \sim R'$. Then, from additivity on $[a, b] \times [c, d]$ and the definition of u_1 extended, it is easily shown that $u_1(x_1) + u_2(x_2) = u_1(y_1) + u_2(y_2)$. On the other hand, if x and y lie above the $u = 2$ curve we have the situation shown in Figure 5.6. Then, by construction and Q1, $(x \sim y, P \sim P') \Rightarrow Q \sim Q'$. By the Figure 5.5 analysis additivity holds for Q and Q', and it readily follows that $u_1(x_1) + u_2(x_2) = u_1(y_1) + u_2(y_2)$. By analogy, additivity holds in each of the four rectangles that have a boundary in common with $[a, b] \times [c, d]$. By induction, additivity holds for every rectangle (complete or truncated) to the right or left of $[a, b] \times [c, d]$ and above or below $[a, b] \times [c, d]$.

The next step is to show that additivity holds throughout $(X_1 \times [c, d]) \cup ([a, b] \times X_2)$. There are no unusual difficulties in this and we omit the proof. It can then be shown that additivity holds in each of the four rectangles that

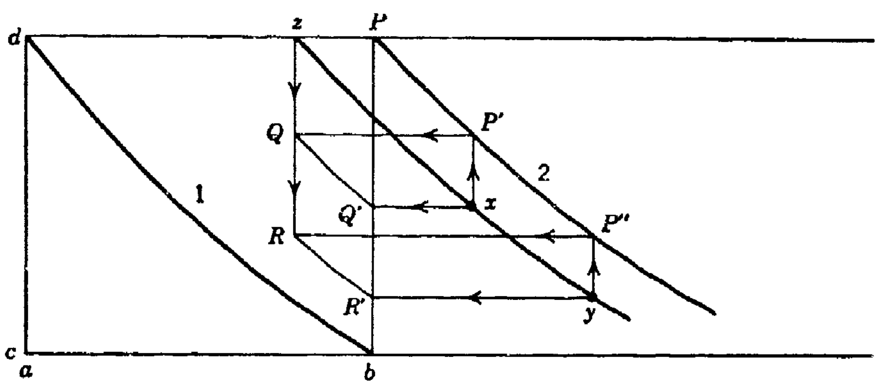

Figure 5.5

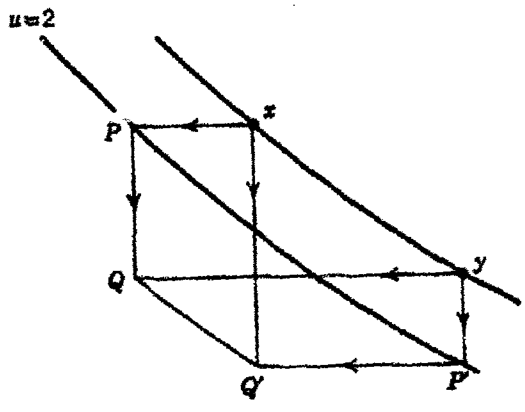

Figure 5.6

have one corner in common with $[a, b] \times [c, d]$ and then that additivity holds on all of $(X_1 \times [c, d]) \cup ([a, b] \times X_2) \cup$ (four corner rectangles). The systematic introduction of new rectangles completes the proof.

Proof, Part II. We now see how the general situation for Theorem 5.4 can be transformed into the structure assumed in Part I of the proof. The hypotheses of the theorem are assumed to hold.

1. By Lemmas 5.3 and 5.1 there is a real-valued function w on $X_1 \times X_2$ that is continuous in $\mathcal{T}_1 \times \mathcal{T}_2$ and satisfies

$$x \prec y \Leftrightarrow w(x) < w(y), \qquad \text{for all } x, y \in X. \tag{5.22}$$

With $(a, b) \in X_1 \times X_2$ fixed let $w_1(x_1) = w(x_1, b)$ and $w_2(x_2) = w(a, x_2)$ for all $x_1 \in X_1$, $x_2 \in X_2$. By Lemma 5.4 and Exercise 3.16, w_i is continuous in $\mathcal{T}_i$ and $W_i = \{w_i(x_i) : x_i \in X_i\}$ is a nondegenerate interval in Re. Let $\mathcal{R}_i$ be the relative usual topology on W_i. Each $(W_i, \mathcal{R}_i)$ is a connected and separable topological space.

2. Let v on $W_1 \times W_2$ be defined by $v(w_1(x_1), w_2(x_2)) = w(x_1, x_2)$. From step 1 it follows that v is well defined and increases in both components. Defining $\prec^*$ on $W_1 \times W_2$ by

$$(c, d) \prec^* (e, f) \Leftrightarrow v(c, d) < v(e, f) \tag{5.23}$$

it follows from (5.22) that

$$(w_1(x_1), w_2(x_2)) \prec^* (w_1(y_1), w_2(y_2)) \Leftrightarrow (x_1, x_2) \prec (y_1, y_2). \tag{5.24}$$

Hence $\prec^*$ is a weak order and it satisfies $(c, d) < (e, f) \Rightarrow (c, d) \prec^* (e, f)$, similar to (5.18). It remains to show that $Q1$ and $Q3$ hold for $\prec^*$ on $W_1 \times W_2$.

3. For $Q1$ suppose for $W_1 \times W_2$ that $(c^1, c^2, c^3) E_3 (d^1, d^2, d^3)$ and $(c^1 \precsim^* d^1, c^2 \precsim^* d^2)$. We need to obtain $d^3 \precsim^* c^3$. Let (x_1^j, x_2^j) for $j = 1, 2, 3$ satisfy $(c_1^j, c_2^j) = (w_1(x_1^j), w_2(x_2^j))$. Define y_i^1, y_i^2, y_i^3 equal to x_i^1, x_i^2, x_i^3 according to the permutations (for $i = 1$ then $i = 2$) that establish $(c^1, c^2, c^3) E_3 (d^1, d^2, d^3)$. Then $(x^1, x^2, x^3) E_3 (y^1, y^2, y^3)$ and by (5.24) and $Q1$ for $<$ on $X_1 \times X_2$, $y^3 \precsim x^3$. Hence, again by (5.24), $d^3 \precsim^* c^3$.

4. To establish $Q3$ for $<^*$ we note first that v is continuous in W_1 and in W_2. For the W_1 proof let $W_1(x_2) = \{w(x_1, x_2): x_1 \in X_1\}$ for each $x_2 \in X_2$, so that $W_1(x_2)$ is an interval for each x_2. By Definition 5.2 we are to show that $\{c: c \in W_1, v(c, d) \in A\} \in \mathfrak{R}_1$ when $d \in W_2$ and $A \in \mathcal{U}$. Let $x_2 \in X_2$ satisfy $w_2(x_2) = d$. Then $\{c: c \in W_1, v(c, d) \in A\} = \{w_1(x_1): x_1 \in X_1, w_1(x_1, x_2) \in A\} = \{w(x_1, b): x_1 \in X_1, \ w(x_1, x_2) \in A \cap W_1(x_2)\}$. Since $w(x_1, x_2) < w(x_1', x_2) \Leftrightarrow w(x_1, b) < w(x_1', b)$, it follows from the continuity of w that if $A \cap W_1(x_2)$ is an open interval in $W_1(x_2)$ then $\{w_1(x_1, b): x_1 \in X_1, w(x_1, x_2) \in A \cap W_1(x_2)\}$ is an open interval in W_1 and hence that $\{c: c \in W_1, v(c, d) \in A\} \in \mathfrak{R}_1$. Thus, if $A \in \mathcal{U}$, then in general $\{c: c \in W_1, \ v(c, d) \in A\} \in \mathfrak{R}_1$. (See Exercise 19.) Hence v is continuous in W_1. The proof for W_2 is similar.

Now suppose $(c, d) <^* (e, f)$. Then $v(c, d) < v(e, f)$ by (5.23). Since v increases and is continuous in each component, it is easily seen that there are intervals $R_1(c)$, $R_1(e) \in \mathfrak{R}_1$ and $R_2(d)$, $R_2(f) \in \mathfrak{R}_2$ such that $(c, d) \in R_1(c) \times R_2(d)$, $(e, f) \in R_1(e) \times R_2(f)$, $t <^* (e, f)$ for all $t \in R_1(c) \times R_2(d)$ and $(c, d) <^* t$ for all $t \in R_1(e) \times R_2(f)$. This is condition 2 of Theorem 3.5 in the $<^*$ context. It then follows from that theorem that $\{r: r \in W_1 \times W_2, r <^* t\} \in \mathfrak{R}_1 \times \mathfrak{R}_2$ and $\{r: r \in W_1 \times W_2, \ t <^* r\} \in \mathfrak{R}_1 \times \mathfrak{R}_2$ for each $t \in W_1 \times W_2$, which is $Q3$ for $<^*$.

Thus, all the hypotheses of Part I hold for $<^*$ on $W_1 \times W_2$, so that there are real-valued continuous functions v_1 on W_1 and v_2 on W_2 that satisfy $(c, d) <^* (e, f) \Leftrightarrow v_1(c) + v_2(d) < v_1(e) + v_2(f)$ and are unique up to similar positive linear transformations when $v_1 + v_2$ additivity holds. Defining $u_i(x_i) = v_i(w_i(x_i))$ we then get $(x_1, x_2) < (y_1, y_2) \Leftrightarrow u_1(x_1) + u_2(x_2) < u_1(y_1) + u_2(y_2)$. Because w_i on X_i is continuous and v_i on W_i is continuous, u_i on X_i is continuous (Exercise 16). ◆

Three or More Factors

Provided that at least three factors actively influence preferences, or are *essential* to use Debreu's term, Debreu's additivity theory with $n \geq 3$ requires only the $m = 2$ part of condition C in Theorem 4.1. For ready comparison with Theorems 5.4 and 5.3 we state his theorem as follows.

THEOREM 5.5. *Suppose* $X = \prod_{i=1}^n X_i$, $n \geq 3$, $<$ *on* X *is a weak order,* $x < y$ *for some* $x, y \in X$ *that differ only in the* ith *components* ($i = 1, \ldots, n$), *and the following hold throughout* X:

*Q1**. $[(x, z)\, E_2\, (y, w),\, x \prec y \text{ or } x \sim y] \Rightarrow \text{not } z \prec w$.

*Q2**. $(X_i, \mathcal{T}_i)$ *is a connected and separable topological space for* $i = 1, \ldots, n$.

*Q3**. $\{x : x \in X, x \prec y\} \in \prod_{i=1}^{n} \mathcal{T}_i$ *and* $\{x : x \in X, y \prec x\} \in \prod_{i=1}^{n} \mathcal{T}_i$.

Then there are real-valued functions $u_1, \ldots, u_n$ *on* $X_1, \ldots, X_n$ *respectively that satisfy* (5.8), *and* $u_1, \ldots, u_n$ *satisfying* (5.8) *are continuous in* $\mathcal{T}_1, \ldots, \mathcal{T}_n$ *respectively and are unique up to similar positive linear transformations.*

Proof, Part I. As for the preceding theorem we consider first the case where each X_i is a nondegenerate real interval, $\mathcal{T}_i$ is the relative usual topology for X_i, and (5.18) holds ($x < y \Rightarrow x \prec y$) along with the other hypotheses of Theorem 5.5.

1. For the same reasons given in Step 1, Part I of the preceding proof, and by Lemma 5.4, there is a continuous (in $\prod \mathcal{T}_i$) $\prec$ – preserving real-valued function v on X that is continuous also in any combination of factors. Moreover, (5.19) holds.

2. Following Debreu (pp. 22–24) we consider first an additive representation for $X_1 \times X_2$. With $a_i \in X_i$ on the interior of X_i for $i > 2$, let

$$H = X_1 \times X_2 \times \{a_3\} \times \cdots \times \{a_n\}$$

and let $\prec^0$ on $X_1 \times X_2$ be the weak order induced by the restriction of $\prec$ on H. By $Q1^*$, $\prec^0$ is independent of the particular $a_3, \ldots, a_n$ values used. Moreover, the conditions in the first paragraph of the preceding Part I proof apply to $\prec^0$ on $X_1 \times X_2$: Q_3 follows easily from Theorem 3.5, but $Q1$ (the C_3 condition) is more difficult to verify.

3. Because of continuity and $x < y \Rightarrow x \prec y$, the former Part I proof used only the indifference part of $Q1$ in the two forms shown in Figure 5.7. Form I was used to establish additivity on $[a, b] \times [c, d]$: Form II was used in extending additivity to all of $X_1 \times X_2$. In either case the $\sim$ part of $Q1$ says

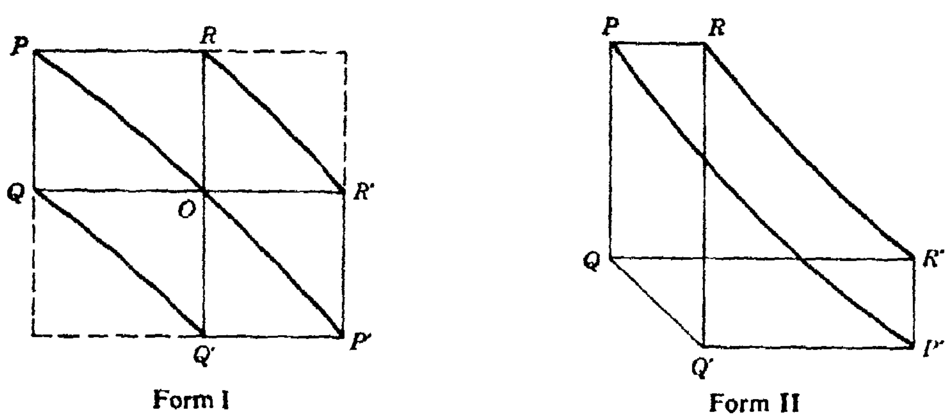

Figure 5.7

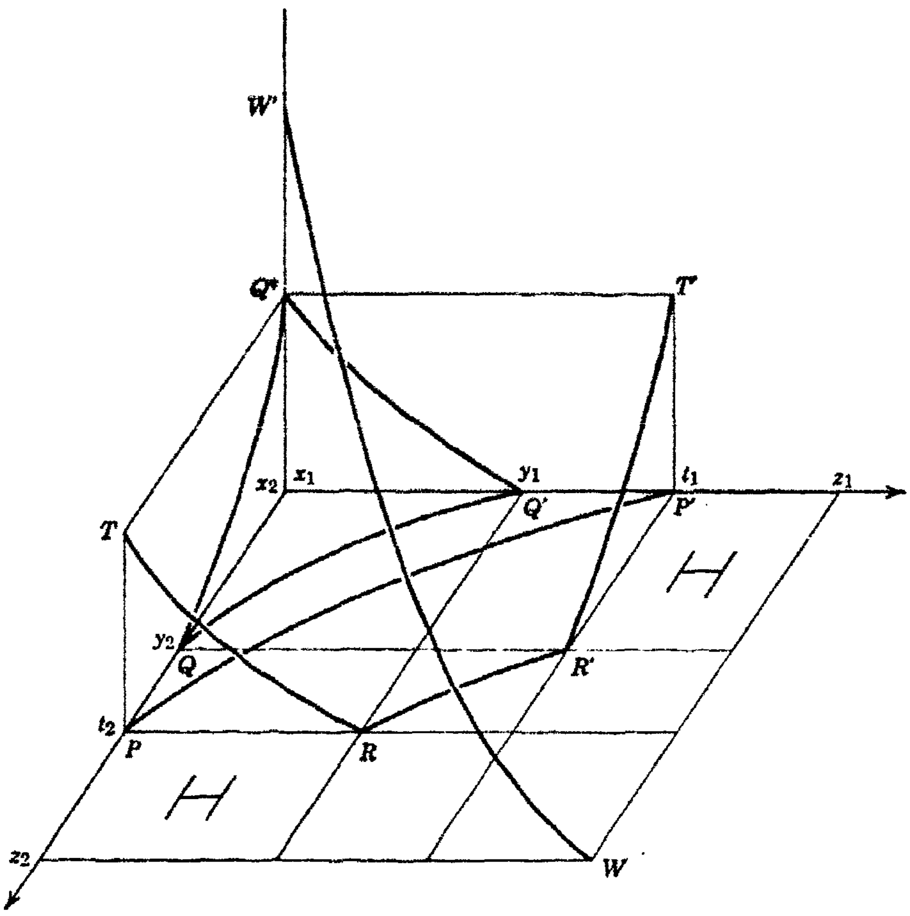

Figure 5.8

that $(P \sim P', Q \sim Q') \Rightarrow R \sim R'$ and $(Q \sim Q', R \sim R') \Rightarrow P \sim P'$. To show that these hold on H we show first that they hold for sufficiently small rectangles in H.

4. Let $x_1 < z_1$ and $x_2 < z_2$ with the differences $z_1 - x_1$ and $z_2 - x_2$ sufficiently small so that there will be a point $W' = (x_1, x_2, b_3, \ldots, b_n) \in X$ that is indifferent to $W = (z_1, z_2, a_3, \ldots, a_n) \in H$. This is shown on Figure 5.8 and follows from continuity and the fact that the a_i were chosen on the interiors of the X_i. Let $y_i, t_i \in X_i$ for $i = 1, 2$ be such that $x_i < y_i < t_i < z_i$ and such that $Q \sim Q'$ and $P \sim P'$. Because $W \sim W'$ there is a $Q^* = (x_1, x_2, c_3, \ldots, c_n)$ in the indifference set (hypersurface) containing Q and Q'. Let T, T', R, and R' be positioned as indicated. Then, by $Q1^*$, $[(Q^*, R)\, E_2\, (Q', T), Q^* \sim Q'] \Rightarrow R \sim T$, $[(Q^*, R')\, E_2\, (Q, T'), Q^* \sim Q] \Rightarrow R' \sim T'$, and $[(P, T')\, E_2\, (P', T), P \sim P'] \Rightarrow T \sim T'$, so that $R \sim R'$ by transitivity. By a similar analysis (take $R \sim R'$, then position P, P'), we have $(Q \sim Q', R \sim R') \Rightarrow P \sim P'$.

5. Suppose $P \sim P'$ and $Q \sim Q'$ as in Form I, Figure 5.7. By repeating the procedure used for positioning the new point P'' in Figure 5.3 we obtain a succession of such points and their associated indifference curves that proceed toward the lower left corner of the rectangle that has P and P' (Figure 5.7) as two corner points. Using the construction procedure of Figure 5.3 at each step, the rectangle is divided into many small rectangles. After some sufficiently large number of steps, the $W \sim W'$ condition of step 5 above will apply to each 2×2 block of four small rectangles, and hence $Q1$ holds in these cases. Beginning in the lower left corner and using the $Q1$ condition on the 2×2 blocks, one can show that, for each small rectangle, the lower right corner is indifferent to the upper left corner. Using transitivity, this leads to $R \sim R'$. Similarly, if $R \sim R'$ and $Q \sim Q'$, we find (by working into the middle from the lower left and upper right corners) that $P \sim P'$. It then follows from the former Part I proof that additive utilities hold in the rectangle with corners P and P' in Form II of Figure 5.7 and from this it follows that $Q1$ holds for Form II. Thus $Q1$ holds in general for H.

6. We know that additive utilities exist for H. Proceeding by induction assume that for each i from 1 to $k - 1$ (≥ 2) there is a continuous, increasing real-valued function u_i on X_i such that the indifference hypersurfaces in $\prod_{i=1}^{k-1} X_i$ (i.e., $\prod_{i=1}^{k-1} X_i \times \prod_{i=k}^{n} \{a_i\}$) are represented by $\sum_{i=1}^{k-1} u_i(x_i) =$ constant. Following Debreu (p. 24) we extend additivity to $\prod_{i=1}^{k} X_i$.

It follows from $Q1^*$ and step 1 that $\sum_{i=1}^{k-1} u_i(x_i) = \sum_{i=1}^{k-1} u_i(y_i) \Leftrightarrow v(x_1, \ldots, x_{k-1}, x_k, a_{k+1}, \ldots, a_n) = v(y_1, \ldots, y_{k-1}, x_k, a_{k+1}, \ldots, a_n)$. Hence, we can define a real-valued function f on $\{\sum_{i=1}^{k-1} u_i(x_i) : x_i \in X_i \text{ for } i = 1, \ldots, k-1\} \times X_k$ by

$$f(\alpha, x_k) = v(x_1, \ldots, x_k, a_{k+1}, \ldots, a_n) \quad \text{when} \quad \sum_1^{k-1} u_i(x_i) = \alpha \qquad \text{for some } x_i.$$

The f increases in each component and is continuous since v is continuous. Let $\Omega = \{v(x_1, \ldots, x_k, a_{k+1}, \ldots, a_n) : x_i \in X_i \text{ for } i = 1, \ldots, k\}$, a real interval. With $\omega \in \Omega$, the set of all (α, x_k) pairs that satisfy

$$f(\alpha, x_k) = \omega \tag{5.25}$$

represents an indifference hypersurface in $\prod_{i=1}^{k} X_i$. Clearly, given $(x_k, \omega) \in X_k \times \Omega$, if (5.25) holds for some $\alpha = \sum_1^{k-1} u_i(x_i)$, this α is unique; we shall call it $g(x_k, \omega)$. It follows that the ω indifference hypersurface represented by (5.25) can be thought of also as the set of all $(x_1, \ldots, x_k)$ for which

$$\sum_{i=1}^{k-1} u_i(x_i) = g(x_k, \omega). \tag{5.26}$$

Let G, a subset of $X_k \times \Omega$, be the domain of definition of g. With $\mathcal{T}_\Omega$ the relative usual topology for Ω, the applicable topology for G is $\mathcal{T}_G = \{G \cap A : A \in \mathcal{T}_k \times \mathcal{T}_\Omega\}$. g is continuous in $\mathcal{T}_G$. (See Exercise 22.)

7. Let (a_k, ω^0) be in the interior of G, and take $(u_1(a_1), \ldots, u_{k-1}(a_{k-1}))$ from the interior of $\{(u_1(x_1), \ldots, u_{k-1}(x_{k-1})): \sum_1^{k-1} u_i(x_i) = g(a_k, \omega^0)\}$:

$$\sum_1^{k-2} u_i(a_i) + u_{k-1}(a_{k-1}) = g(a_k, \omega^0). \tag{5.27}$$

Next, let $(x_k, \omega) \in G$ be near enough to (a_k, ω^0) so that the operations used with (5.28) and (5.29) are possible. Select $(c_1, \ldots, c_{k-1}) \in \prod_1^{k-1} X_i$ for which

$$\sum_1^{k-2} u_i(a_i) + u_{k-1}(c_{k-1}) = g(x_k, \omega^0) \tag{5.28}$$

$$\sum_1^{k-2} u_i(c_i) + u_{k-1}(a_{k-1}) = g(a_k, \omega). \tag{5.29}$$

By (5.27) and (5.28), $(a_1, \ldots, a_{k-2}, a_{k-1}, a_k) \sim (a_1, \ldots, a_{k-2}, c_{k-1}, x_k)$ since both are on the ω^0 indifference hypersurface. Then, by $Q1^*$, $(c_1, \ldots, c_{k-2}, a_{k-1}, a_k) \sim (c_1, \ldots, c_{k-2}, c_{k-1}, x_k)$. Since the first of these is on the ω indifference hypersurface by (5.29), so is the latter:

$$\sum_1^{k-2} u_i(c_i) + u_{k-1}(c_{k-1}) = g(x_k, \omega). \tag{5.30}$$

Subtracting (5.27) from (5.28) and (5.29) from (5.30) we get

$$g(x_k, \omega) = g(a_k, \omega) + g(x_k, \omega^0) - g(a_k, \omega^0). \tag{5.31}$$

8. Let V be a rectangle in G whose sides are parallel to the axes (of X_k and Ω) and which contains (a_k, ω^0) and permits the operations used on (5.28) and (5.29) for each $(x_k, \omega) \in V$. By (5.31) and Lemma 5.4, g on V can be written as the sum of increasing continuous functions of ω and x_k, say

$$g(x_k, \omega) = h(\omega) - u_k(x_k). \tag{5.32}$$

This analysis applies to each (a_k, ω^0) in the interior of G: each such (a_k, ω^0) will have an associated V rectangle in G within which g can be decomposed as in (5.32). Suppose $V \cap V' \neq \varnothing$ with

$$g(x_k, \omega) = h(\omega) - u_k(x_k) \qquad \text{for } (x_k, \omega) \in V \tag{5.33}$$

$$g(x_k, \omega) = h'(\omega) - u_k'(x_k) \qquad \text{for } (x_k, \omega) \in V'. \tag{5.34}$$

Fix $(b, \omega^0) \in V \cap V'$ and transform h' and u_k' by adding constants so that

$$h'(\omega^0) = h(\omega^0) \quad \text{and} \quad u_k'(b) = u_k(b). \tag{5.35}$$

Suppose $(x_k, \omega) \in V \cap V'$. Then, by the parallel sides condition, (b, ω) and (x_k, ω^0) are in $V \cap V'$. Hence, using (5.33) and (5.34),

$$u_k(x_k) = -g(x_k, \omega^0) + h(\omega^0)$$
$$u_k'(x_k) = -g(x_k, \omega^0) + h'(\omega^0)$$

so that $u_k(x_k) = u'_k(x_k)$ on using (5.35). Similarly, $h(\omega) = h'(\omega)$. Thus, under the alignment of (5.35), $(h, u_k) = (h', u'_k)$ on $V \cap V'$. It follows that h and u_k can be defined so as to satisfy (5.32) throughout the interior of G. Continuity then insures that (5.32) holds on all of G.

9. Substitution of (5.32) into (5.26) yields $\sum_{i=1}^{k} u_i(x_i) = h(\omega)$ as a representation of the ω indifference hypersurface in $\prod_{i=1}^{k} X_i$. By induction, each indifference hypersurface in $\prod_{i=1}^{n} X_i$ can be represented by $\sum_{i=1}^{n} u_i(x_i) =$ constant, with each u_i continuous and increasing in x_i.

Proof, Part II. The proof that the general situation of Theorem 5.5 can be transformed into the structure of Part I in this proof is similar to the Part II proof of Theorem 5.4. ◆

5.5 SUMMARY

Although a very general theory of additivity has been developed by Tversky (1967), it becomes somewhat complex and difficult to interpret in an easy way when infinite sets are involved. The reader interested in a very general theory should consult this paper.

When rather strong structural conditions, such as weak order, $X = \prod_{i=1}^{n} X_i$, solvability, and so forth are assumed to hold, less general but more easily interpreted additivity theories result. One of these, developed by Luce and Tukey (1964) and Luce (1966), is algebraic in nature and involves the assumption that differences in the levels of some factors can be offset (in the preference sense) by compensating differences in the levels of other factors. As shown by Luce (1966) it is possible to weaken this unrestricted solvability condition and still obtain results similar to those in Section 5.2. The theory that results from restricted solvability is very similar to the topological additivity theory of Debreu (1960) as reviewed in Section 5.4. In all the theories noted in this paragraph, the independence condition C_3 of Theorem 4.1 is sufficient for additivity, but $C_2(m = 2)$ can be used when there are more than two factors because C_3 then follows from C_2 and the other conditions. In these well-structured theories additive utilities are unique up to similar positive linear transformations, and in Debreu's theory each u_i on X_i is continuous in the topology associated with $(<, X_i)$.

INDEX TO EXERCISES

1–3. Lexicographic orders and additivity. 4. $mx + nx = (m + n)x$. 5–6. Strictly ordered groups. 7. Commutative group. 8. Similar positive linear transformations. 9. Unbounded utilities. 10. Countable sets applicability. 11–12. Closure. 13. Product topology. 14–15.

Products, intersections, unions. 16–22. Continuity. 23–24. Insufficiency of $C_2(P1^*, Q1^*)$ when $n = 2$. 25. Mean-variance criterion for normal probability distributions.

Exercises

1. For each of the following cases $X = X_1 \times X_2$ and $(x_1, x_2) \prec (y_1, y_2) \Leftrightarrow (x_1, x_2) <^L (y_1, y_2) \Leftrightarrow x_1 < y_1$ or $(x_1 = y_1, x_2 < y_2)$. Verify the assertions made.
a. $X_1 = \{0, 1\}$, $X_2 = \{r : r$ is a rational number$\}$. Additive utilities exist.
b. $X_1 = \{r : r$ is a rational number$\}$, $X_2 = \{0, 1\}$. Additive utilities don't exist.
c. $X_1 = X_2 = \{j : j$ is an integer$\}$. Additive utilities exist.

2. (*Continuation.*) Even though additive utilities exist in Exercise 1*c*, $(Y, +, <^L)$ as defined preceding Theorem 5.1 for this case is a non-Archimedean strictly ordered group and therefore there is no f on Y that satisfies (5.3) and (5.4). Discuss this situation further.

3. Let $X = X_1 \times X_2 \times X_3$ with $X_1 = \{0, 1\}$, $X_2 = \{1, 2, \ldots\}$.
a. If $X_3 = \{0, 1\}$ prove that additive utilities do not exist when $x \prec y \Leftrightarrow x <^L y$.
b. If $X_3 = \mathrm{Re}$ and $x \prec y \Leftrightarrow x <^L y$ show that there is a countable subset of X that is $\prec$-order dense in X.

4. Let $(Y, +)$ be a group. Show that if $m > 0$ and $n < 0$ are integers then $mx + nx = (m + n)x$ whenever $x \in Y$.

5. Prove that a strictly ordered group is Archimedean when (5.3) and (5.4) hold.

6. With L_x and U_x as defined in the proof of Theorem 5.1, prove that there is a unique real number $g(x)$ such that $m/n \le g(x) \le r/s$ for all $m/n \in L_x$ and $r/s \in U_x$.

7. Let $Y = \{0, a\}$, $a \ne 0$, and define $ma = 0$ when m is an even integer and $ma = a$ when m is an odd integer. Define $+$ fully so that $(Y, +)$ is a commutative group.

8. Show that additive utilities are unique up to similar positive linear transformations when $X = \{x_1, y_1\} \times \{x_2, y_2\}$ and $(x_1, x_2) \prec (x_1, y_2) \sim (y_1, x_2) \prec (y_1, y_2)$. (Assume that $\prec$ is a weak order.)

9. Verify that under the hypotheses of Theorem 5.2 u_1 and u_2 in (5.6) must be unbounded when $x \prec y$ for some $x, y \in X$.

10. When $x \prec y$ for some $x, y \in X$, can X_1 and X_2 be countable under the hypotheses of Theorem 5.2? Is the same thing true for the hypotheses of Theorem 5.4?

11. Let $X = \mathrm{Re}$ with the usual topology $\mathfrak{U}$. Specify the closure of (*a*) $\{0, 1, 2, \ldots\}$; (*b*) $\{r : 0 < r < 1$ and r is rational$\}$; (*c*) $\{1/n : n = 1, 2, 3, \ldots\}$; (*d*) $\{m/2^n : m = 0, 1, \ldots, 2^n, n = 1, 2, \ldots\}$.

12. Let X be all rational points in Re, with the relative usual topology $\{X \cap A : A \in \mathfrak{U}\}$. What is the closure of X?

13. Let $(X_i, \mathfrak{T}_i)$ be a topological space for $i = 1, 2, \ldots, n$, and let $\prod^* \mathfrak{T}_i$ be the

family of sets formable by arbitrary unions of the sets in $\{\prod_{i=1}^{n} A_i : A_i \in \mathcal{T}_i$ for $i = 1, \ldots, n\}$. Prove that $\prod^* \mathcal{T}_i = \prod \mathcal{T}_i$.

14. Let $X = \prod_{i=1}^{n} X_i$, $A_i^j \subseteq X_i$ for $j = 1, \ldots, m$ and $i = 1, \ldots, n$. Prove that $\prod_{i=1}^{n} (\bigcap_{j=1}^{m} A_i^j) = \bigcap_{j=1}^{m} (\prod_{i=1}^{n} A_i^j)$.

15. (*Continuation.*) Let $X = \prod_{i=1}^{n} X_i$, $A_i(t) \subseteq X_i$ for all $t \in T$, where T is an arbitrary set. Verify

a. $\prod_{i=1}^{n} (\bigcap_{t\in T} A_i(t)) = \bigcap_{t\in T} (\prod_{i=1}^{n} A_i(t))$,

b. $\bigcup_{t\in T} (\prod_{i=1}^{n} A_i(t)) \subseteq \prod_{i=1}^{n} (\bigcup_{t\in T} A_i(t))$,

c. Show by example that $\bigcup_T (\prod A_i(t))$ can be a proper subset of $\prod (\bigcup_T A_i(t))$.

16. Let $(X, \mathcal{R})$, $(Y, \mathcal{S})$, and $(Z, \mathcal{T})$ be topological spaces and suppose f on X into Y is $\mathcal{R} - \mathcal{S}$ continuous and g on Y into Z is $\mathcal{S} - \mathcal{T}$ continuous. Let $h(x) = g(f(x))$ for all $x \in X$. Prove that h is $\mathcal{R} - \mathcal{T}$ continuous.

17. Using the first part of the proof of Lemma 5.4 as a guide, prove that if $(X, \mathcal{R})$, $(Y, \mathcal{S})$, and $(Z, \mathcal{T})$ are topological spaces, if f on X into Y is $\mathcal{R} - \mathcal{S}$ continuous and if g on X into Z is $\mathcal{R} - \mathcal{T}$ continuous, then h on X into $Y \times Z$, defined by $h(x) = (f(x), g(x))$, is $\mathcal{R} - (\mathcal{S} \times \mathcal{T})$ continuous.

18. With $\mathcal{U}$ the usual topology for Re, verify that $A \in \mathcal{U}$ if and only if A is the union of open intervals in Re.

19. (*Continuation.*) Argue that a real function f on X is continuous in the topology $\mathcal{T}$ for X if and only if A is an open interval in Re implies that $f^{-1}(A) = \{x : x \in X, f(x) \in A\}$ is in $\mathcal{T}$.

20. With $(X, \mathcal{R})$ and $(Y, \mathcal{S})$ topological spaces and f a function on X into Y, let $f(X) = \{y : y \in Y$ and $y = f(x)$ for some $x \in X\}$. Show that if f is $\mathcal{R} - \mathcal{S}$ continuous, then f is $\mathcal{R} - \{f(X) \cap S : S \in \mathcal{S}\}$ continuous. Thus, a continuous function is continuous also with respect to the relative topology for its range.

21. Let f be a real, strictly increasing (or strictly decreasing) function on a real interval $[a, b]$, and suppose that the range of f, $f(X)$, is a real interval. Prove that f is continuous.

22. Suppose X, Y, and Z are real intervals, f on $X \times Y$ onto Z is strictly increasing in each variable and is continuous. For each $(y, z) \in Y \times Z$ for which there is an $x \in X$ that satisfies $f(x, y) = z$, let $g(y, z)$ equal x when $f(x, y) = z$. Let $G \subseteq Y \times Z$ be the domain of g. Prove that g is continuous.

23. Given $X = [1, \infty) \times [1, \infty)$ and

$$u(x_1, x_2) = x_1 x_2 + x_1^{x_2} \qquad \text{for each} \quad (x_1, x_2) \in X,$$

suppose $(x_1, x_2) \prec (y_1, y_2)$ if and only if $u(x_1, x_2) < u(y_1, y_2)$, for all (x_1, x_2), $(y_1, y_2) \in X$. Verify that all the hypotheses of Theorem 5.4 hold except for $Q1$, and that $Q1^*$ holds for this case (i.e., C_2). Do additive utilities exist in this case? Why not?

24. (*Continuation:* due to David Krantz.) Given $X = (0, \infty) \times (0, \infty)$ and

$$u(x_1, x_2) = \begin{cases} x_1 x_2 + x_1^{x_2} & \text{if } 1 \le x_1, 1 \le x_2 \\ x_1(x_2 + 1) & \text{if } 0 < x_1 \le 1 \le x_2 \\ 2x_1 x_2 & \text{if } 0 < x_1, 0 < x_2 \le 1 \end{cases}$$

suppose $(x_1, x_2) \prec (y_1, y_2)$ if and only if $u(x_1, x_2) < u(y_1, y_2)$, for all (x_1, x_2), $(y_1, y_2) \in X$. Verify that all hypotheses of Theorems 5.2 and 5.4 hold with the exception of *P*1 or *Q*1, and that *P*1 fails but *P*1* or *Q*1* holds.

25. Let X be the set of all normal probability distributions on the real line. Each such distribution is completely known when its mean μ and standard deviation σ (≥ 0) are specified, so that we can represent X by the set X' of all ordered pairs (μ, σ) for which μ is a real number and $\sigma \geq 0$. If the hypotheses of Theorem 5.4 hold for $\prec$ on X' then there are continuous real-valued functions f on $(-\infty, \infty)$ and g on $[0, \infty)$ such that, for every (μ, σ) and (μ^*, σ^*) in X',

$$(\mu, \sigma) \prec (\mu^*, \sigma^*) \Leftrightarrow f(\mu) + g(\sigma) < f(\mu^*) + g(\sigma^*).$$

If you are familiar with normal probability distributions, comment on the reasonableness of the hypotheses (in particular *Q*1) in the case where each normal distribution represents a course of action that is a gamble for amounts of money.

Chapter 6

COMPARISON OF PREFERENCE DIFFERENCES

All preference axioms in preceding chapters and those in Parts II and III involve only simple preference comparisons ($<$). In this chapter, however, we shall consider a "strength-of-preference" notion that involves comparisons of preference differences. We will use a binary relation $<^*$ on pairs of ordered pairs in $X \times X$.

We interpret $(x, y) <^* (z, w)$ to mean that the degree of preference for x over y is less than the degree of preference for z over w. The "degree of preference" for x over y can of course be "negative" if y is preferred to x.

For conceptual clarity I shall use $x - y$ to denote an ordered pair $(x, y) \in X \times X: x - y = (x, y)$. Thus, $x - y <^* z - w$ will be used in place of, and is identical to, $(x, y) <^* (z, w)$. This notation suggests some conditions that may clarify the notion of directed preference difference comparisons, such as

$$x - y <^* z - w \Rightarrow w - z <^* y - x, \tag{6.1}$$

$$x - y <^* z - w \Rightarrow x - z <^* y - w. \tag{6.2}$$

In our utility representations $x - y <^* z - w$ will be associated with $u(x) - u(y) < u(z) - u(w)$. In distinction to this approach Suppes and Winet (1955) work with undirected or absolute difference comparisons and associate the preference degree between x and y with $|u(x) - u(y)|$. They use also a simple preference relation ($<$). With directed differences $<$ can be defined directly from $<^*$, such as

$$x < y \Leftrightarrow x - x <^* y - x, \tag{6.3}$$

but at least one author, Armstrong (1939), has taken issue with this. His idea, which is not in vogue today, was to take $<^*$ as a precisely "measurable" notion so that, for example, if $x < y$ then, by gradual changes, one can always find a z between x and y so that $z - x \sim^* y - z$, and eventually

obtain $x - y <^* z - w \Leftrightarrow u(x) - u(y) < u(z) - u(w)$: but at the same time he championed intransitive indifference for $\sim$ with $x < y$ only if the difference $u(y) - u(x)$ exceeds a minimal positive threshold value.

6.1 "MEASURABLE" UTILITY

Before we look at some formal theory, other remarks should be made. Defining

$$x - y \sim^* z - w \Leftrightarrow (\text{not } x - y <^* z - w, \text{ not } z - w <^* x - y) \quad (6.4)$$

it seems clear, Armstrong (1939) and others to the contrary notwithstanding, that there is no more (and probably less) reason to suppose that $\sim^*$ is transitive than to suppose that $\sim$ is transitive. For example, can you find one and only one value of x for which $\$x - \$0 \sim^* \$100000 - \$x$? If you can, I venture to say that your discriminatory judgment is rather more acute than that of most mortals.

Since its introduction by Pareto (1927, p. 16) and Frisch (1926), the idea of comparable preference differences has been severely criticized, and for reasons that go deeper than the discriminatory vagueness that may lead to intransitive $\sim^*$. One charge has been that the notion has no operational meaning. Because of this, several "operational" modes for making comparisons have been suggested, including the following three, where we assume for convenience that $y < x < w < z$.

1. To compare $x - y$ and $z - w$, compare a 50-50 gamble resulting in either x or w with a 50-50 gamble resulting in either y or z. If the former is preferred, take $z - w <^* x - y$, and so forth.
2. To compare $x - y$ and $z - w$ imagine that you already have y and w and can either exchange y for x or exchange w for z. If you prefer the former exchange take $z - w <^* x - y$, and so forth.
3. Assuming that x, y, z, and w are nonmonetary, estimate the minimum bonus $\$a$ for which $x \sim y + \$a$, and estimate the minimum bonus $\$b$ for which $z \sim w + \$b$. If $\$a < \$b$ take $x - y <^* z - w$.

Of these three we must reject the second since it violates the hypothesis that X is a set of mutually exclusive alternatives, in which case it makes little if any sense to suppose that you already have both y *and* w. The third approach, which might appeal to some people, is suspect first for the reason that it presupposes a form of independence between X and the monetary bonuses (as in a two-factor situation in Chapters 4 and 5) and second that, even if independence applies, there is some question about defining a strength-of-preference notion on the basis of simple preference comparisons.

This last clause applies also, as noted by Weldon (1950) and Ellsberg (1954), among others, to the 50-50 gambles device. Simple comparisons between even-chance gambles as a basis for defining degree of preference seem to distort the notion introduced by Pareto and Frisch. Included in this distortion is the addition of chance, which plays no part in the basic notion. Along with Weldon and Ellsberg, I would have no quarrel with an individual who judges that $\$30 - \$0 <^* \$100 - \$40$ but prefers an even-chance gamble between \$30 and \$40 to one between \$0 and \$100. The latter judgment involves the individual's attitude toward taking chances, an attitude we feel is not part of the $<^*$ notion.

If we do in fact reject such approaches we may be driven back to the idea of the early writers on this subject, that $<^*$ comparisons are essentially a matter of direct self-interrogation as to whether your degree of preference for x over y exceeds, equals, or is less than your degree of preference for z over w. As noted above, this is rejected by some because of its "nonoperational" character.

Others dislike the idea of direct preference-difference comparisons for the reason that, under sufficiently powerful conditions on $<^*$, one must logically accept the ability to "measure" preference differences introspectively much as one would go about measuring lengths with a measuring rod. This implication of the "measurability" of utility has caused much commotion in the literature: some writers who accept the concept of simple preference comparisons find it impossible to endorse the notion of "measurable" utility. Pareto, in fact, denounced the very notion he introduced when he found that it was not needed to derive certain results in the theory of static, riskless, consumer demand. On the other hand, Frisch (1964) remains an advocate of "measurable" utility: in the cited paper, on the subject of dynamic (time-dependent) consumer demand theory, he points out that several attractive results cannot be obtained without some notion of "measurable" utility.

For some people, the direct, introspective "measurability" pill may be easier to swallow when intransitive $\sim^*$ is allowed to enter the theory. Although our preference-difference comparisons may not be as precise as length comparisons made with precision instruments, I do not feel that this is sufficient reason to abandon the idea of such comparisons.

6.2 THEORY WITH FINITE SETS

Using the method of Adams (1965), we now state and prove two representation theorems for preference-difference comparisons when X is finite. Both are incorporated in Theorem 6.1. The $A \Leftrightarrow A^*$ theorem permits intransitive $\sim^*$, but the $B \Leftrightarrow B^*$ theorem takes $\sim^*$ as transitive. The A theorem is

proved by Adams (1965). An equivalent of the B theorem is proved by Scott (1964).

THEOREM 6.1. *Suppose X is finite. Then*

A. $[x^1, \ldots, x^m, w^1, \ldots, w^m$ *is a permutation of* $y^1, \ldots, y^m, z^1, \ldots, z^m$ *and* $x^j - y^j <^* z^j - w^j$ *for all* $j < m] \Rightarrow$ *not* $x^m - y^m <^* z^m - w^m$;

B. $[x^1, \ldots, x^m, w^1, \ldots, w^m$ *is a permutation of* $y^1, \ldots, y^m, z^1, \ldots, z^m$ *and* $x^j - y^j <^* z^j - w^j$ *or* $x^j - y^j \sim^* z^j - w^j$ *for each* $j < m] \Rightarrow$ *not* $x^m - y^m <^* z^m - w^m$;

for all $x^j, y^j, z^j, w^j \in X$ *and* $m = 2, 3, \ldots$, *if and only if there is a real-valued function u on X such that, for all* $x, y, z, w \in X$,

A^*. $x - y <^* z - w \Rightarrow u(x) - u(y) < u(z) - u(w)$;
B^*. $x - y <^* z - w \Leftrightarrow u(x) - u(y) < u(z) - u(w)$.

It is easily seen that $A^* \Rightarrow A$ and $B^* \Rightarrow B$. A does not require $<^*$ to be transitive although the transitive closure of $<^*$ under A is a strict partial order. A does not imply either (6.1) or (6.2). On the other hand, B implies that $<^*$ is a weak order along with (6.1) and (6.2). For asymmetry, B says that $x - y <^* z - w \Rightarrow$ not $z - w <^* x - y$, since x, z, w, y is a permutation of y, w, z, x. Negative transitivity then follows from B: (not $x - y <^* z - w$, not $z - w <^* r - s) \Rightarrow (z - w \leqslant^* x - y,\ r - s \leqslant^* z - w) \Rightarrow$ not $x - y <^* r - s$. With $<$ as defined in (6.3), B implies that $<$ on X is a weak order. Here and later, $\leqslant^* = <^* \cup \sim^*$.

Sufficiency Proofs. Let A hold. To apply the Theorem of The Alternative (Theorem 4.2) let $c = (u(t^1), u(t^2), \ldots, u(t^N))$ where $X = \{t^1, \ldots, t^N\}$. Let $\mathcal{A}$ be the set of all $x - y <^* z - w$ statements. If $\mathcal{A} = \varnothing$, A^* is immediate. If $\mathcal{A} \neq \varnothing$, each corresponding $u(x) - u(y) < u(z) - u(w)$ translates into a $c \cdot a^k > 0$ statement, which gives a system like (4.4). If this system has no c solution then, by Theorem 4.2 and the fact that the $a_j^k \in \{-1, 0, 1\}$ for all j and k, there are non-negative integers r_k at least one of which is positive such that $\sum_k r_k a_j^k = 0$ for $j = 1, \ldots, N$. From the original $x - y <^* z - w$ statements it then follows that there is a sequence $x^1 - y^1 <^* z^1 - w^1, \ldots, x^m - y^m <^* z^m - w^m$ with $x^1, \ldots, x^m, w^1, \ldots, w^m$ a permutation of $y^1, \ldots, y^m, z^1, \ldots, z^m$. If $m > 1$, this violates A. If $m = 1$, it yields $x - y <^* x - y$ or else $x - x <^* y - y$, each of which violates A. Hence there is a c solution.

Let B hold. Axiom B implies as a special case that if (in the two-dimensional sense) $((x^1, y^1), \ldots, (x^m, y^m))\ E_m\ ((z^1, w^1), \ldots, (z^m, w^m))$ and if $x^j - y^j <^* z^j - w^j$ or $x^j - y^j \sim^* z^j - w^j$ for each $j < m$, then not $x^m - y^m <^* z^m - w^m$. It follows immediately from Theorem 4.1C that there

are real-valued functions u_1 and u_2 on X such that $x - y <^* z - w \Leftrightarrow u_1(x) + u_2(y) < u_1(z) + u_2(w)$. Also, B implies that $x - y <^* z - w \Leftrightarrow w - z <^* y - x$. Hence $x - y <^* z - w \Leftrightarrow u_1(w) + u_2(z) < u_1(y) + u_2(x)$. Defining $u(x) = u_1(x) - u_2(x)$ it then follows that $x - y <^* z - w \Leftrightarrow u(x) - u(y) < u(z) - u(w)$. ◆

6.3 REVIEW OF INFINITE-SET THEORIES

In this section we review some theories that assume that $<^*$ on $X \times X$ is a weak order and imply that there is a real-valued function u on X satisfying

$$x - y <^* z - w \Leftrightarrow u(x) - u(y) < u(z) - u(w), \qquad \text{for all } x, y, z, w \in X \quad (6.5)$$

that is "unique up to a positive linear transformation." This means that if u satisfies (6.5) then v satisfies (6.5) also if and only if there are real numbers $a > 0$ and b such that

$$v(x) = au(x) + b, \qquad \text{for all } x \in X. \quad (6.6)$$

The two-factor additivity theories of Chapter 5 can be adapted to the present case. Suppose, for example, that there are real-valued functions u_1 and u_2 on X such that

$$x - y <^* z - w \Leftrightarrow u_1(x) + u_2(y) < u_1(z) + u_2(w), \qquad \text{for all } x, y, z, w \in X, \quad (6.7)$$

with u_1 and u_2 unique up to similar positive linear transformations. Suppose also that (6.1) holds. Then, as in the proof of Theorem 6.1B, u on X, defined by $u(x) = u_1(x) - u_2(x)$, satisfies (6.5). In addition, u is unique up to a positive linear transformation. For suppose that u and v satisfy (6.5). Defining $u_1(x) = u(x)$, $u_2(x) = -u(x)$, $v_1(x) = v(x)$ and $v_2(x) = -v(x)$, it follows from (6.5) that (6.7) holds for (u_1, u_2) and for (v_1, v_2). Since v_1 is a positive linear transformation of u_1, v is a positive linear transformation of u.

From this reasoning and Theorem 5.4, the following axioms, after Debreu (1960), imply a u for (6.5) that is continuous in $\mathcal{T}$:

*A*1. $x - y <^* z - w \Rightarrow w - z <^* y - x$,

*A*2. $[((x^1, y^1), (x^2, y^2), (x^3, y^3))\, E_3\, ((z^1, w^1), (z^2, w^2), (z^3, w^3)),\ x^j - y^j <^* z^j - w^j$ *or* $x^j - y^j \sim^* z^j - w^j$ *for* $j = 1, 2] \Rightarrow$ *not* $x^3 - y^3 <^* z^3 - w^3$,

*A*3. $(X, \mathcal{T})$ *is a connected and separable topological space,*

*A*4. $\{x - y : x - y \in X \times X, x - y <^* z - w\} \in \mathcal{T} \times \mathcal{T}$ *and* $\{x - y : x - y \in X \times X, z - w <^* x - y\} \in \mathcal{T} \times \mathcal{T}$, *for every* $z - w \in X \times X$.

Algebraic Axioms

Suppes and Winet (1955), Scott and Suppes (1958), and Suppes and Zinnes (1963, pp. 34–38) present nontopological axioms that imply a u for (6.5) that

is unique up to a positive linear transformation. The first four Suppes-Zinnes axioms are equivalent to $B1$ and $B2$:

$B1$. *$<^*$ on $X \times X$ is a weak order,*
$B2$. (6.1) *and* (6.2).

Their final three axioms, rather than using the complete $A2$, are based on algebraic conditions. With $<$ on X and $\sim^*$ on $X \times X$ as defined in (6.3) and (6.4), $x - yM^1z - w$ means that $x - y \sim^* z - w$ and $y \sim z$. (That is, the preference interval from y to x "equals" the preference interval from w to z and the two intervals are contiguous.) Proceeding recursively, $x - yM^{n+1}z - w$ means that there are $s, t \in X$ such that $x - yM^n s - t$ and $s - tM^1z - w$. The final three axioms are: for every $x, y, z, w \in X$,

$B3$. *$x - s \sim^* s - y$ for some $s \in X$,*
$B4$. *$(y < x, z - w <^* x - y) \Rightarrow (y < s < x, z - w \leqslant^* x - s)$ for some $s \in X$,*
$B5$. *$(y < x, x - y \leqslant^* z - w) \Rightarrow (z - sM^n t - w, z - s \leqslant^* x - y)$ for some $s, t \in X$ and some positive integer n.*

$B3$ is the midpoint or bisection axiom, similar to Armstrong's notion following (6.3). In nontrivial cases, $B3$ requires X to be infinite. $B4$ is like a continuity condition, and $B5$ is a structural-Archimedean axiom. $B5$ says that if the difference $x - y$ is "positive" then, no matter how large $z - w$ happens to be, there is an n such that the $z - w$ interval can be divided into $n + 1$ equal parts no one of which is larger than $x - y$.

Pfanzagl's Theory

Pfanzagl (1959) presents axioms that, under one interpretation, imply (6.5) with u unique up to a positive linear transformation. His general theory uses a set X that is connected (topologically) and a function f on $X \times X$ into X. Instead of $<^*$ he uses $<$ along with f. However, in the interpretation of this chapter, $<^*$ is not completely absent since $f(x, y)$ is interpreted as a point in X that is midway in preference between x and y, like s in $B3$.

In addition to a continuity axiom, Pfanzagl's theory uses the following assumptions:

$C1$. *$<$ on X is a weak order,*
$C2$. *$x < y \Rightarrow f(x, z) < f(y, z)$ and $f(z, x) < f(z, y)$ for every $z \in X$; $x \sim y \Rightarrow f(x, z) \sim f(y, z)$ and $f(z, x) \sim f(z, y)$ for every $z \in X$,*
$C3$. $f(f(x, y), f(z, w)) \sim f(f(x, z), f(y, w))$.

$C3$ is the *bisymmetry axiom.* These axioms (including continuity) imply that

there is a real-valued function u on X that satisfies

$$x < y \Leftrightarrow u(x) < u(y) \tag{6.8}$$

$$u(f(x, y)) = pu(x) + qu(y) + r \tag{6.9}$$

for all $x, y \in X$ and is unique up to a positive linear transformation.

Under the interpretation of f as a midpoint function, two more axioms arise:

$C4$. $f(x, x) \sim x$
$C5$. $f(x, y) \sim f(y, x)$.

When $x < y$ for some $x, y \in X$, $C4$ and $C5$ require $p = q = \frac{1}{2}$ and $r = 0$ in (6.9). It follows that $f(x, y) < f(z, w) \Leftrightarrow u(x) + u(y) < u(z) + u(w)$. (6.5) then follows when $<^*$ on $X \times X$ is defined as follows:

$$x - y <^* z - w \Leftrightarrow f(x, w) < f(z, y). \tag{6.10}$$

6.4 SUMMARY

The notion of comparable preference differences is (with the exception of Exercise 17) the only strength-of-preference or preference intensity concept that appears in this book. The additive utility theories of Chapters 4 and 5, although mathematically similar to the theories in this chapter, are based solely on simple preference comparisons and involve no higher-order preference concepts.

With $x - y <^* z - w$ interpreted as "your degree of preference for z over w exceeds your degree of preference for x over y," the conditions that relate $x - y <^* z - w$ to $u(x) - u(y) < u(z) - u(w)$ are similar to the conditions used in two-factor additivity theories. Exceptions to this arise in (6.1) and (6.2), which are addressed specifically to the preference-difference notion and have no counterparts in preceding chapters.

INDEX TO EXERCISES

1–3. Even-chance gambles theory. 4–6. (6.1) and (6.2). 7–9. (6.3). 10. Condition B. 11. Semiordered preference differences. 12–13. Algebraic theory. 14–16. Pfanzagl's conditions. 17–18. "Twice as happy."

Exercises

1. Interpret $(x, y) \prec (z, w)$ to mean that a 50-50 gamble between $z, w \in X$ is preferred to a 50-50 gamble between $x, y \in X$. Assuming that X is finite, give

necessary and sufficient conditions for $\prec$ on $X \times X$ for each of the following two utility representations: (*a*) $(x, y) \prec (z, w) \Rightarrow u(x) + u(y) < u(z) + u(w)$; (*b*) $(x, y) \prec (z, w) \Leftrightarrow u(x) + u(y) < u(z) + u(w)$.

2. (*Continuation.*) Using Theorem 5.4 argue that, when $A1$ of Section 6.3 is replaced by $(x, y) \sim (y, x)$ for all $x, y \in X$, and $(\prec^*, \sim^*)$ in $A2$, $A3$, and $A4$ is replaced by $(\prec, \sim)$, then there is a real-valued function u on X that satisfies $(x, y) \prec (z, w) \Leftrightarrow u(x) + u(y) < u(z) + u(w)$ and is continuous in $\mathscr{C}$ and unique up to a positive linear transformation.

3. (*Continuation.*) Interpret $f(x, y)$ in Pfanzagl's theory as an element in X that is indifferent to a 50-50 gamble between x and y. Show that $(x, y) \prec (z, w) \Leftrightarrow u(x) + u(y) < u(z) + u(w)$ follows from (6.8) and (6.9) when $C4$ and $C5$ are used and $(x, y) \prec (z, w) \Leftrightarrow f(x, y) \prec f(z, w)$.

4. Prove that [$\prec^*$ is irreflexive, (6.2)] $\Rightarrow x - x \sim^* y - y$.

5. Prove that [(6.1), (6.2)] $\Rightarrow (x - y \sim^* z - w \Leftrightarrow x - z \sim^* y - w \Leftrightarrow w - z \sim^* y - x)$.

6. Prove that [$\prec^*$ is asymmetric, (6.2), $x - y \sim^* z - w \Leftrightarrow x - z \sim^* y - w$] $\Rightarrow$ (6.1).

7. Suppose that $\prec^*$ on $X \times X$ is a strict partial order, (6.2) holds, $x \prec y$, $y \prec z$, and $x \prec z$ according to (6.3) and, with $a \approx^* b \Leftrightarrow (a \sim^* c \Leftrightarrow b \sim^* c$, for all $c \in X \times X)$, $r - r \approx^* s - s$ for all $r, s \in X$. Show that: (*a*) $x - y \prec^* z - x$; (*b*) $x - y \prec^* z - y$; (*c*) $x - z \prec^* y - z$; (*d*) $y - x \prec^* z - x$; (*e*) $z - y \prec^* z - x$.

8. (*Continuation.*) Show that [$\prec^*$ is a strict partial order, (6.1), (6.2), $x - x \approx^* y - y$ for all $x, y \in X$] $\Rightarrow$ $\prec$ on X as defined by (6.3) is a strict partial order.

9. Show that [$\prec^*$ is a weak order, (6.2), $x - y \sim^* z - w \Leftrightarrow x - z \sim^* y - w$] $\Rightarrow$ $\prec$ on X as defined by (6.3) is a weak order.

10. Show that B_2 of Theorem 6.1 (B with $m = 2$) implies (6.2) and $x - y \sim^* z - w \Leftrightarrow x - z \sim^* y - w$.

11. Prove the following theorem. If X is finite, if $\prec^*$ on $X \times X$ is irreflexive, and if [$x^1, \ldots, x^{2m}, w^1, \ldots, w^{2m}$ is a permutation of $y^1, \ldots, y^{2m}, z^1, \ldots, z^{2m}$, $x^j - y^j \sim^* z^j - w^j$ for $j = 1, \ldots, m$, $x^j - y^j \prec^* z^j - w^j$ for $j = m + 1, \ldots, 2m - 1$] $\Rightarrow$ not $x^{2m} - y^{2m} \prec^* z^{2m} - w^{2m}$, for all positive integers m and $x^j, y^j, z^j, w^j \in X$, then there is a real-valued function u on X such that

$$x - y \prec^* z - w \Leftrightarrow u(x) - u(y) + 1 < u(z) - u(w), \qquad \text{for all} \quad x, y, z, w \in X.$$

12. Interpret M^1, M^2, and M^3 (Section 6.3) in terms of points on a line.

13. Show that $B1$ and $B2$ in Section 6.3 imply that if $x - y \preccurlyeq^* y - y$ and $y - z \preccurlyeq^* w - t$ then $x - z \preccurlyeq^* w - t$. (Use Exercises 4 and 5. This exercise is due to Michael Levine: see Suppes and Zinnes (1963, p. 35).)

14. Show that [(6.8), (6.9), $C4$, $C5$, $x \prec y$ for some $x, y \in X$] $\Rightarrow p = q = \frac{1}{2}$, $r = 0$.

15. With $\prec^*$ defined from $\prec$ on X as in (6.10), prove the following.

a. ($C1$, $C2$, $C3$) $\Rightarrow$ $\prec^*$ on $X \times X$ is a weak order. (Due to Luce and Tukey (1964, p. 14).)

b. $(C1, C5) \Rightarrow (6.1)$ and (6.2).
c. $(C1, C4) \Rightarrow x - s \sim^* s - y$ for some $s \quad X$.
d. $(C1, C2, C3) \Rightarrow A2$ (in the Debreu axioms).

16. Let condition B_m of Theorem 6.1 hold for $m \leq 6$. Assume also that $f(x, y) = z \Rightarrow x - z \sim^* z - y$ and let (6.10) apply. Prove that $C3$, Pfanzagl's bisymmetry axiom, follows.

17. Galanter (1962) asks the following type of question: What amount of money, as a gift, would make you feel twice as happy as you'd feel if you were to receive a gift of \$10? If the response in \$45 (the median for one sample), it is suggested that we set $u(\$45) = 2u(\$10)$, with $u(\$0) = 0$. This is the same as taking $u(\$45) - u(\$10) = u(\$10) - u(\$0)$ so that \$10 is midway in preference between \$0 and \$45. Do you feel that this midpoint interpretation is reasonable in view of the question that gave rise to it and the strength-of-preference interpretation used in the chapter?

18. (*Continuation.*) A motorist is asked for his reaction to delays at toll booths with the question: What waiting time r would make you twice as mad as you would be if you had to wait for time t? Given the set of (t, r) pairs $\{(1, 3), (3, 8), (8, 18), (18, 30), (30, 45), (45, 60), (60, 75), (75, 90)\}$ and taking $u(r) - u(0) = 2[u(t) - u(0)]$ for each of the eight (t, r) pairs, set $u(0) = 0$ and $u(1) = -1$ and sketch u on $[0, 100]$.

Chapter 7

PREFERENCES ON HOMOGENEOUS PRODUCT SETS

A homogeneous product set has the form $X = A \times A \times \cdots$. If A is repeated n times, we write $X = A^n$. A common interpretation for A^n is that there are n time periods and $(x_1, \ldots, x_n) \in A^n$ represents a series of similar events that can be selected or occur during the n periods: x_i is the event for period i. $(x_1, \ldots, x_n)$ could be a series of annual incomes for the next n years or, in a single-period context, x_i could be the amount of money allocated to the ith of n activities.

With $X = A^n$, this chapter examines concepts for the time context, including persistence, impatience, and discounting. Our usage of these terms is based on the work of Koopmans (1960), Koopmans, Diamond and Williamson (1964), and Diamond (1965) in a denumerable-period formulation.

Throughout, $\prec$ on X will be assumed to be a weak order. Since the independence notions of Chapters 4 and 5 are relevant for $X = A^n$, we shall consider, in conjunction with the foregoing concepts, special cases of

$$x \prec y \Leftrightarrow \sum_{i=1}^{n} u_i(x_i) < \sum_{i=1}^{n} u_i(y_i), \qquad \text{for all } x, y \in A^n. \tag{7.1}$$

One such case is the no time preference situation where ρ is a real-valued function on A and

$$x \prec y \Leftrightarrow \sum_{i=1}^{n} \rho(x_i) < \sum_{i=1}^{n} \rho(y_i), \qquad \text{for all } x, y \in A^n. \tag{7.2}$$

Given (7.1), it is easily shown that there is a ρ that satisfies (7.2) if and only if $(x_1, \ldots, x_n) \sim (y_1, \ldots, y_n)$ whenever $x_1, \ldots, x_n$ is a permutation of $y_1, \ldots, y_n$. In the time context this says that times of occurrence of various events have no affect on preferences, which is often false. Somewhat more realistic special cases of (7.1) will be considered later.

7.1 PERSISTENCE AND IMPATIENCE

Two notions that postulate forms of regularity of preferences in the homogeneous time context are persistence and impatience. Persistence applies when similar preferences hold in the various periods. Impatience says that you'd rather have more preferred things happen sooner than later. In the following definitions $\bar{a}$ denotes the constant alternative that yields $a \in A$ in every time period: $\bar{a} = (a, \ldots, a)$.

Definition 7.1. $<$ on A^n is *persistent* if and only if $(x_1, \ldots, x_{i-1}, a, x_{i+1}, \ldots, x_n) < (x_1, \ldots, x_{i-1}, b, x_{i+1}, \ldots, x_n) \Rightarrow (y_1, \ldots, y_{j-1}, a, y_{j+1}, \ldots, y_n) < (y_1, \ldots, y_{j-1}, b, y_{j+1}, \ldots, y_n)$ whenever $i, j \in \{1, \ldots, n\}$ and all four n-tuples are in A^n. $<$ on A^n is *impatient* if and only if $\bar{a} < \bar{b} \Rightarrow (x_1, \ldots, x_{i-1}, a, b, x_{i+2}, \ldots, x_n) < (x_1, \ldots, x_{i-1}, b, a, x_{i+2}, \ldots, x_n)$ and $\bar{a} \sim \bar{b} \Rightarrow (x_1, \ldots, x_{i-1}, a, b, x_{i+2}, \ldots, x_n) \sim (x_1, \ldots, x_{i-1}, b, a, x_{i+2}, \ldots, x_n)$ whenever $i \in \{1, \ldots, n-1\}$ and the n-tuples are in A^n.

Persistence seems reasonable when the n-tuples in X represent income streams over a period of n years. Impatience might also hold in this case. The reverse of impatience could hold in some situations for people who prefer to postpone favorable events, perhaps to increase their anticipatory pleasure or for a variety of other reasons. The reverse of persistence might arise from a desire for variety, as in the chicken-steak example preceding Section 4.1.

When $<$ is a weak order, $<$ is persistent implies that $<_i$ on A, defined by $a <_i b \Leftrightarrow (x_1, \ldots, x_{i-1}, a, x_{i+1}, \ldots, x_n) < (x_1, \ldots, x_{i-1}, b, x_{i+1}, \ldots, x_n)$ for some $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n \in A$, is a weak order (which also follows from condition $C_2(m = 2)$, Theorem 4.1) and that $<_1, \ldots, <_n$ are identical (which does not follow from C_2).

In our definition of impatience, a and b are in contiguous time periods. A more general case of impatience arises when

$$\bar{a} < (\sim)\bar{b} \Rightarrow (x_1, \ldots, x_{i-1}, a, x_{i+1}, \ldots, x_{j-1}, b, x_{j+1}, \ldots, x_n)$$
$$< (\sim)(x_1, \ldots, x_{i-1}, b, x_{i+1}, \ldots, x_{j-1}, a, x_{j+1}, \ldots, x_n) \quad (7.3)$$

for any $1 \leq i < j \leq n$ and $x_1, \ldots, x_n \in A$. This does not follow from persistence and impatience. The following theorem amplifies this statement.

THEOREM 7.1. ($<$ *is a persistent and impatient weak order on* A^n) *does not imply* (7.3). ($<$ *is an impatient weak order on* A^n *that satisfies condition* C_2 *of Theorem* 4.1) *implies* (7.3).

Proof. For the latter assertion it suffices to show that the hypotheses imply that $(a, x_2, \ldots, x_{n-1}, b) < (\sim)(b, x_2, \ldots, x_{n-1}, a)$ when $\bar{a} < (\sim)\bar{b}$.

Given $\bar{a} < (\sim)\bar{b}$, repeated applications of impatience give $(a, b, \ldots, b) < (\sim)(b, a, b, \ldots, b) < (\sim)(b, b, a, b, \ldots, b) < (\sim) \cdots < (\sim)(b, \ldots, b, a)$ so that $(a, b, \ldots, b) < (\sim)(b, \ldots, b, a)$. Since $((a, b, \ldots, b), (b, x_2, \ldots, x_{n-1}, a))\ E_2\ ((b, \ldots, b, a), (a, x_2, \ldots, x_{n-1}, b))$, $(a, x_2, \ldots, x_{n-1}, b) < (\sim)(b, x_2, \ldots, x_{n-1}, a)$ follows from condition C with $m = 2$.

To verify the negative assertion, take $A = \{a, b, c\}$, $n = 3$, and let $<$ on A^3 be defined by (7.1) when the u_i on A take the values shown in the following array. Clearly, $<$ is persistent and as we shall note in the parentheses it is

	u_1	u_2	u_3
a	0	0	0
b	10	9	8
c	20	15	12

impatient $(8 < 9 < 10, 12 < 15 < 20, 10 + 15 < 20 + 9, 9 + 12 < 15 + 8)$ and in fact satisfies (7.3). In the middle of $<$ we find $\cdots < (b, a, c) < (a, c, b) < (b, c, a) < (b, b, b) \sim (a, c, c) < \cdots$. Let $<' = <$ except that we replace $(a, c, b) < (b, c, a)$ by $(a, c, b) \sim' (b, c, a)$. With this one change, persistence and impatience hold also for $<'$, but (7.3) fails since $\bar{a} <' \bar{b}$. ◆

Additive Utilities

When (7.1) holds and $<$ is persistent, each u_i function has the same order on A, as illustrated with $A = [0, 1]$ on the left of Figure 7.1. When $<$ is impatient also we get a picture like that on the right of the figure in which

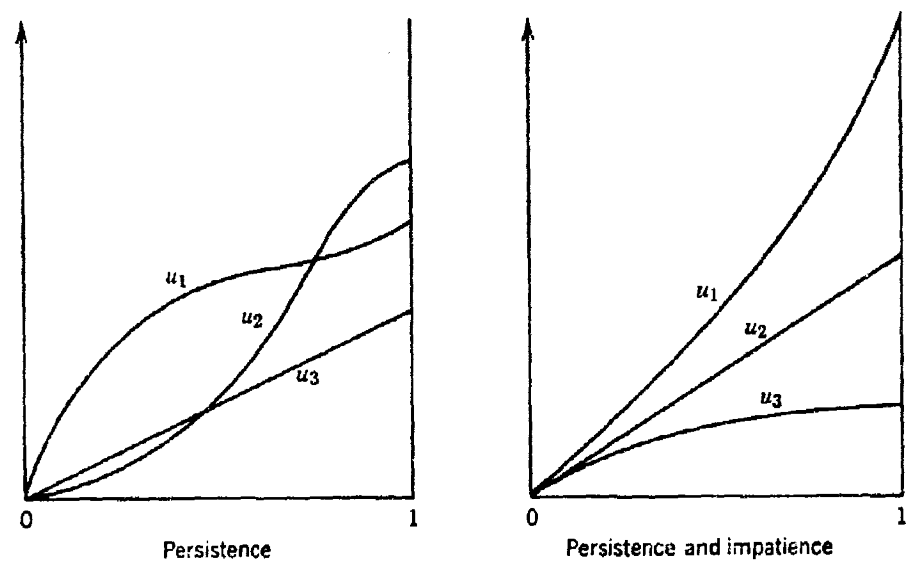

Figure 7.1 Additive utilities on $[0, 1]^3$.

$u_1(b) - u_1(a) > u_2(b) - u_2(a) > u_3(b) - u_3(a)$ whenever $b > a$ (i.e., $\bar{a} < \bar{b}$), which says that the vertical distance between u_1 and u_2, and between u_2 and u_3, increases as b increases.

Additive utilities can of course hold when impatience holds and persistence fails. For, with $A = \{a, b\}$ and $n = 2$, it is easily seen that $(a, b) < (a, a) < (b, b) < (b, a)$ has a u_1 and u_2 that satisfy (7.1). Since $\bar{a} < \bar{b}$ and $(a, b) < (b, a)$, $<$ is impatient. However, $<$ is not persistent since $(a, b) < (b, b)$ and $(b, b) < (b, a)$.

7.2 PERSISTENT PREFERENCE DIFFERENCES

We shall now look at a higher-order persistence notion based on the degree of preference relation $<^*$ on $X \times X$ used in Chapter 6, along with the weak-order difference representation

$$x - y <^* z - w \Leftrightarrow u(x) - u(y) < u(z) - u(w), \qquad \text{for all } x, y, z, w \in A^n. \tag{7.4}$$

As in Definition 7.1, $\bar{a} = (a, \ldots, a)$ in the following.

Definition 7.2. $<^*$ on $A^n \times A^n$ is *persistent* if and only if

$$x - y <^* z - w \Leftrightarrow \bar{x}_j - \bar{y}_j <^* \bar{z}_j - \bar{w}_j \tag{7.5}$$

whenever $j \in \{1, \ldots, n\}$, $x_i = y_i$ and $z_i = w_i$ for all $i \neq j$, and $x, y, z, w \in A^n$.

This says that the order of preference differences with constant alternatives dictates the order of differences for each j, other things being equal. With $n = 2$, $<^*$ is persistent implies that if $(a, x_2) - (b, x_2) <^* (c, y_2) - (d, y_2)$ for some $x_2, y_2 \in A$ then this holds for every $x_2, y_2 \in A$ and, in addition, $(x_1, a) - (x_1, b) <^* (y_1, c) - (y_1, d)$ holds for every $x_1, y_1 \in A$.

Part of the power of persistent $<^*$ is shown by the next theorem.

THEOREM 7.2. *If u on A^n satisfies* (7.4) *and if $<^*$ is persistent then there are real-valued functions $u_1, \ldots, u_n$ on A for which*

$$u(x_1, \ldots, x_n) = \sum_{i=1}^{n} u_i(x_i), \qquad \textit{for all } x \in A^n, \tag{7.6}$$

and, for every $a, b, c, d \in A$ and $i, j \in \{1, \ldots, n\}$.

$$u_i(a) - u_i(b) < u_i(c) - u_i(d) \Leftrightarrow u_j(a) - u_j(b) < u_j(c) - u_j(d). \tag{7.7}$$

When $<$ is defined as in (6.3), (7.1) follows immediately from (7.6) and (7.4). Hence, additive utilities exist for A^n when (7.4) holds and $<^*$ is persistent.

Proof. Let u satisfy (7.4) and assume that $<^*$ is persistent. Fix $e \in A$, assign $u_1(e), \ldots, u_n(e)$ so that $u(\bar{e}) = \sum u_i(e)$, and define u_i on A, for $i = 1, \ldots, n$, by

$$u_i(a) = u(e, \ldots, e, a, e, \ldots, e) - \sum_{j \neq i} u_j(e), \qquad \text{for all } a \in A. \tag{7.8}$$

To verify (7.6), let $\alpha^i = (x_1, \ldots, x_{i-1}, x_i, e, \ldots, e)$, $\beta^i = (x_1, \ldots, x_{i-1}, e, e, \ldots, e)$, and $\gamma^i = (e, \ldots, e, x_i, e, \ldots, e)$, for $2 \leq i \leq n$. If $\alpha^i - \beta^i <^* \gamma^i - \bar{e}$ then $\bar{x}_i - \bar{e} <^* \bar{x}_i - \bar{e}$ by (7.5), and similarly if $\gamma^i - \bar{e} <^* \alpha^i - \beta^i$. Hence $\alpha^i - \beta^i \sim^* \gamma^i - \bar{e}$ so that, by (7.4),

$$\begin{aligned} u(x_1, \ldots, x_i, e, \ldots, e) - u(x_1, \ldots, x_{i-1}, e, \ldots, e) \\ = u(e, \ldots, e, x_i, e, \ldots, e) - u(\bar{e}). \end{aligned}$$

Summing from $i = 2$ to $i = n$ and using $u(\bar{e}) = \sum u_i(e)$ and (7.8) we get

$$u(x_1, \ldots, x_n) - u(x_1, e, \ldots, e) = \sum_{i=2}^{n} u_i(x_i) - \sum_{i=2}^{n} u_i(e),$$

which yields (7.6) after $u(x_1, e, \ldots, e)$ is transposed and (7.8) is used again. (7.7) follows easily from (7.4), (7.5) and (7.6). ◆

Weighted Additivity

In the rest of this section we shall consider a form of weighted, additive utilities that is less general than (7.1) and more general than (7.2). This is the form

$$x < y \Leftrightarrow \sum_{i=1}^{n} \lambda_i \rho(x_i) < \sum_{i=1}^{n} \lambda_i \rho(y_i), \qquad \text{for all } x, y \in A^n, \tag{7.9}$$

where $\lambda_i > 0$ for each i and ρ is a real-valued function on A. It is easily seen that, when (7.1) holds, (7.9) can hold if and only if there are u_i satisfying (7.1) that are pairwise related by positive linear transformations, say with $u_j = a_j u_1 + b_j$ and $a_j > 0$ for $j = 2, \ldots, n$.

In the time context the λ_i are weights for the different periods. If $\lambda_1 > \lambda_2 > \cdots > \lambda_n$, we could call them discount factors: $\lambda_1 > \cdots > \lambda_n$ follows from (7.9) when $<$ is impatient and $x < y$ for some $x, y \in X$. If $\lambda_1 < \cdots < \lambda_n$, the λ_i might be referred to as markup factors.

In general, (7.1) along with $<$ persistent is insufficient for (7.9). As this is written, I do not know of any set of axioms for $<$ on A^n that, even when A is finite, is necessary and sufficient for (7.9). For this reason, and because (7.9) implies (7.7) when $u_i = \lambda_i \rho$, we shall consider a pathway to (7.9) that leads through (7.4) and makes the assumption that $<^*$ is persistent. Even here we shall note a negative conclusion before giving sufficient conditions for (7.9).

THEOREM 7.3. *Suppose* (7.4) *holds and* $<^*$ *is persistent. Then with* $<$ *defined by* $x < y \Leftrightarrow x - x <^* y - x$, *there may not exist* $\lambda_i > 0$ *and a real-valued function* ρ *on* A *that satisfy* (7.9). *This conclusion holds even when* u *in* (7.4) *is unique up to a positive linear transformation.*

Proof. Let $A = \{a, b, c\}$ and $n = 3$, and let (7.1) hold with the u_i defined as follows:

	a	b	c
u_1	0	1	3
u_2	0	2	5
u_3	0	3	9

Define u by (7.6) and take $x - y <^* z - w \Leftrightarrow u(x) - u(y) < u(z) - u(w)$. Because $u(\bar{b}) - u(\bar{a}) < u(\bar{c}) - u(\bar{b})$ and $u_i(b) - u_i(a) < u_i(c) - u_i(b)$ for $i = 1, 2, 3$, $<^*$ is persistent. In defining λ_1, λ_2, λ_3 and ρ for (7.9) we can, with no loss in generality, set $\lambda_1 = 1$, $\rho(a) = 0$ and $\rho(b) = 1$. Then, since $(c, a, a) \sim (a, a, b) \sim (b, b, a)$, $\rho(c) = \lambda_3 = \lambda_2 + 1$. This along with $(a, a, c) \sim (b, c, b)$ gives $(\lambda_2 + 1)^2 = 1 + \lambda_2(\lambda_2 + 1) + (\lambda_2 + 1)$ according to (7.9), and this reduces to $1 = 2$, which is false. Hence (7.9) cannot hold. Moreover, u is unique up to a positive linear transformation when it satisfies (7.4). This follows from the fact that each of the 25 other $u(x_1, x_2, x_3)$ can be written solely in terms of $u(a, a, a)$ and $u(b, a, a)$ when (7.4) holds. ◆

Sufficient Conditions for Weighted Additivity

Despite Theorem 7.3 there are axioms implying (7.4) which imply (7.9) also when $<^*$ is persistent. We consider one such case, based on Debreu's theory. The following correspond to $A1$–$A4$ in Section 6.3. $X = A^n$.

$A1'$. $x - y <^* z - w \Rightarrow w - z <^* y - x$,

$A2'$. *If* x^1, x^2, x^3 *is a permutation of* z^1, z^2, z^3, *and* y^1, y^2, y^3 *is a permutation of* w^1, w^2, w^3, *and if* $x^j - y^j <^* z^j - w^j$ *or* $x^j - y^j \sim^* z^j - w^j$ *for* $j = 1, 2$, *then not* $x^3 - y^3 <^* z^3 - w^3$,

$A3'$. $(A, \mathcal{T})$ *is a connected and separable topological space*,

$A4'$. $\{x - y : x - y \in X \times X, x - y <^* z - w\} \in \mathcal{T}^{2n}$ *and* $\{x - y : x - y \in X \times X, z - w <^* x - y\} \in \mathcal{T}^{2n}$ *for every* $z - w \in X \times X$.

$\mathcal{T}^{2n}$ is the product topology for $X \times X = A^n \times A^n$. By Lemma 5.3, $A3'$ says that $(A^n, \mathcal{T}^n)$ and $(X \times X, \mathcal{T}^{2n})$ are connected and separable topological spaces. It then follows from Theorem 5.4 and $A1'$ that there is a continuous (in $\mathcal{T}^n$) real-valued function u on X that satisfies (7.4) and is unique up to a positive linear transformation.

Let u_i on A be defined by (7.8) in the proof of Theorem 7.2. Since u is continuous in $\mathcal{T}^n$, u_i is continuous in $\mathcal{T}$ for each i. Let $<^*$ on $X \times X$ be

persistent and define $<^*$ on $A \times A$ by

$$a - b <^* c - d \Leftrightarrow \bar{a} - \bar{b} <^* \bar{c} - \bar{d}, \qquad \text{for all } a, b, c, d \in A.$$

It then follows from Theorem 7.2 and persistence that

$$a - b <^* c - d \Leftrightarrow u_i(a) - u_i(b) < u_i(c) - u_i(d), \qquad \text{for all } i. \quad (7.10)$$

This is (7.4) in miniature, for A instead of X. Since u_i is continuous and (7.10) holds for each i, the correspondents of $A1'$–$A4'$ hold for $<^*$ on A^2 for each i. It follows that u_i and u_j are related by a positive linear transformation. In particular there are positive $\alpha_2, \ldots, \alpha_n$, and $\beta_2, \ldots, \beta_n$, such that $u_j(a) = \alpha_j u_1(a) + \beta_j$ for all $a \in A$, $j = 2, \ldots, n$. Letting $\rho \equiv u_1$, $\lambda_1 = 1$, $\lambda_j = \alpha_j$ for $j > 1$, (7.6) gives $u(x) = \sum_{i=1}^n \lambda_i \rho(x_i) +$ constant, which, on using (6.3) and (7.4) gives $x < y \Leftrightarrow \sum \lambda_i \rho(x_i) < \sum \lambda_i \rho(y_i)$, all $\lambda_i > 0$. This proves the first part of the following theorem.

THEOREM 7.4. *Suppose $X = A^n$, $<^*$ on $X \times X$ is persistent, and $A1'$, $A2'$, $A3'$, and $A4'$ hold. Then there are $\lambda_i > 0$ and a continuous (in $\mathcal{T}$) real-valued function ρ on A that satisfy* (7.9) *when $<$ on A^n is defined from $<^*$ on $A^n \times A^n$ by $x < y \Leftrightarrow x - x <^* y - x$. If in addition $n > 1$ and $x < y$ for some $x, y \in X$ and if $\lambda_i' > 0$ and ρ' on A satisfy* (7.9) *along with $\lambda_i > 0$ and ρ on A, then there are numbers $\alpha > 0$, $\beta > 0$ and γ such that*

$$\lambda_i' = \alpha\lambda_i \qquad i = 1, \ldots, n \qquad (7.11)$$

$$\rho'(a) = \beta\rho(a) + \gamma \qquad \textit{for all } a \in A. \qquad (7.12)$$

Proof. For the uniqueness assertions take ρ and the λ_i as defined for the first part. Let $u_i(a) = \lambda_i\rho(a)$ and $u(x) = \sum u_i(x_i)$. Then, as in the first part of the proof, u is continuous and hence, by Theorem 3.5, $\{x : x < y\} \in \mathcal{T}^n$ and $\{x : y < x\} \in \mathcal{T}^n$, which establish condition $Q3$ of Theorem 5.4 ($n = 2$) or condition $Q3^*$ of Theorem 5.5 ($n > 2$). With $<^*$ persistent and $x < y$ for some $x, y \in X$, each of the n factors has an active influence on $<$. Since the other conditions of Theorem 5.4 or Theorem 5.5 are easily seen to hold for $<$ on A^n, by (7.9), it follows that the $\lambda_i\rho$ in (7.9) are unique up to similar positive linear transformations. Hence $\lambda_i' > 0$ and ρ' satisfy (7.9) if and only if there is a $k > 0$ and β_i such that $\lambda_i'\rho'(a) = k\lambda_i\rho(a) + \beta_i$ for $i = 1, \ldots, n$. Since this gives $\rho'(a) = (k\lambda_i/\lambda_i')\rho(a) + \beta_i/\lambda_i'$, ρ' is a positive linear transformation of ρ as in (7.12). Also, since ρ is not constant on A ($x < y$ for some x, y), $k\lambda_i/\lambda_i' = k\lambda_j/\lambda_j'$ for all i, j, or $\lambda_j' = (\lambda_1'/\lambda_1)\lambda_j$ for $j = 2, \ldots, n$. Set $\alpha = \lambda_1'/\lambda_1$. (7.11) then follows. ◆

7.3 CONSTANT DISCOUNT RATES

Although persistent preference differences were used to obtain (7.9) for arbitrary positive λ_i, special cases of (7.9) can be derived using only the

simple preference relation $\prec$. One of these is (7.2). Another occurs when $\lambda_{i+1}/\lambda_i = \pi$ for $i = 1, 2, \ldots, n-1$ with $\pi > 0$, in which case (7.9) reduces to

$$x \prec y \Leftrightarrow \sum_{i=1}^{n} \pi^{i-1}\rho(x_i) < \sum_{i=1}^{n} \pi^{i-1}\rho(y_i), \qquad \text{for all } x, y \in A^n. \tag{7.13}$$

If $\pi = 1$ we have (7.2), the case of no time preference. If $\pi < 1$, (7.13) represents the case where utilities are discounted at a constant rate, which arises under (7.13) when $\prec$ is impatient. When $\pi > 1$, utilities are marked up at a constant rate.

One way to obtain (7.13) is to begin with Debreu's additivity theory. Taking $n \geq 3$, we shall use the hypotheses of Theorem 5.5 applied to $X = A^n (X_i = A$ for each $i)$ along with one more condition. The new condition is referred to as temporal consistency by Williams and Nassar (1966) and as stationarity by Koopmans (1960).

Definition 7.3. $\prec$ on A^n is *stationary* if and only if there is an $e \in A$ such that, for all $x_1, \ldots, x_{n-1}, y_1, \ldots, y_{n-1} \in A$,

$$(x_1, \ldots, x_{n-1}, e) \prec (y_1, \ldots, y_{n-1}, e) \Leftrightarrow (e, x_1, \ldots, x_{n-1}) \prec (e, y_1, \ldots, y_{n-1}). \tag{7.14}$$

In going from $(x_1, \ldots, x_{n-1}, e)$ to $(e, x_1, \ldots, x_{n-1})$ each x_i is updated by one period and e is shifted from the last period to the first. Stationarity says that preferences do not change under such shifts.

THEOREM 7.5. *If the hypotheses of Theorem 5.5 hold for $X = A^n$ and if $\prec$ on A^n is stationary then there is a positive number π and a continuous real-valued function ρ on A that satisfy (7.13). Moreover, π is unique and ρ is unique up to a positive linear transformation.*

Proof. Let the hypotheses hold, with continuous u_i for (7.2) unique up to similar positive linear transformations. Define $\prec$ on A^{n-1} by

$$(x_1, \ldots, x_{n-1}) \prec (y_1, \ldots, y_{n-1}) \Leftrightarrow (x_1, \ldots, x_{n-1}, e) \prec (y_1, \ldots, y_{n-1}, e).$$

It follows from (7.2) and (7.14) that, for all $(x_1, \ldots, x_{n-1}), (y_1, \ldots, y_{n-1}) \in A^{n-1}$,

$$(x_1, \ldots, x_{n-1}) \prec (y_1, \ldots, y_{n-1}) \Leftrightarrow \sum_{i=1}^{n-1} u_i(x_i) < \sum_{i=1}^{n-1} u_i(y_i)$$

$$(x_1, \ldots, x_{n-1}) \prec (y_1, \ldots, y_{n-1}) \Leftrightarrow \sum_{i=1}^{n-1} u_{i+1}(x_i) < \sum_{i=1}^{n-1} u_{i+1}(y_i).$$

It follows from these two expressions and Theorems 5.4 and 5.5 that there is a $\pi > 0$ and numbers $\beta_1, \ldots, \beta_{n-1}$ such that

$$u_{i+1}(a) = \pi u_i(a) + \beta_i \qquad \text{for all } a \in A; \qquad i = 1, \ldots, n-1.$$

Using this recursively to express each u_i in terms of u_1 and letting $\rho \equiv u_1$, substitution into (7.2) yields (7.13).

Suppose (7.13) holds then along with $x < y \Leftrightarrow \sum_i \lambda^{i-1}\sigma(x_i) < \sum_i \lambda^{i-1}\sigma(y_i)$. From Debreu's uniqueness up to similar positive linear transformations it follows that there are numbers $\alpha > 0$ and $\beta_1, \ldots, \beta_{n-1}$ such that

$$\lambda^{i-1}\sigma(a) = \alpha\pi^{i-1}\rho(a) + \beta_i \qquad \text{for all } a \in A; \qquad i = 1, \ldots, n-1.$$

With $i = 1$ this gives $\sigma(a) = \alpha\rho(a) + \beta_1$. Substituting for σ with $i > 1$ we then have $\lambda^{i-1}\alpha\rho(a) + \lambda^{i-1}\beta_1 = \pi^{i-1}\alpha\rho(a) + \beta_i$ which, since ρ is not constant on A, requires $\lambda = \pi$. ◆

7.4 SUMMARY

When $X = A^n$ and i indexes time, new concepts come into play, including no time preference, impatience, persistent preferences, persistent preference differences, and stationarity. These concepts can apply whether or not utilities are additive over the n periods.

The most general special case of additivity considered in this chapter is the weighted form $x < y \Leftrightarrow \sum \lambda_i\rho(x_i) < \sum \lambda_i\rho(y_i)$, with $\lambda_i > 0$ for each i. Debreu's topological theory for weak ordered preference differences along with persistent preference differences implies this form. Additive utilities, but not necessarily the weighted form given here, arise from the representation $x - y <^* z - w \Leftrightarrow u(x) - u(y) < u(z) - u(w)$ along with persistent preference differences.

Under appropriately strong axioms for additive utilities based on simple preference comparisons, the form $x < y \Leftrightarrow \sum \pi^{i-1}\rho(x_i) < \sum \pi^{i-1}\rho(y_i)$ can result when $<$ is assumed to be stationary. If $<$ is impatient also then $0 < \pi < 1$.

INDEX TO EXERCISES

1–3. No time preference. 4–5. Persistent preferences. 6. Impatience. 7. Persistent differences. 8. Nonhomogeneous preference difference additivity. 9–10. Weighted additivity. 11–12. Constant discount rate. 13–14. Present monetary value.

Exercises

1. Given (7.1) prove that (7.2) follows when $(x_1, \ldots, x_n) \sim (y_1, \ldots, y_n)$ whenever $x_1, \ldots, x_n$ is a permutation of $y_1, \ldots, y_n$. Define ρ by $\rho(a) = \sum_{i=1}^n u_i(a)$.

2. With $X \subseteq A^n$, let $(x^1, \ldots, x^m)\ E_m^*\ (y^1, \ldots, y^m) \Leftrightarrow [m > 1,\ x^1, \ldots, x^m, y^1, \ldots, y^m \in X$; the number of times $a \in A$ appears as a component in $(x^1, \ldots, x^m)$

equals the number of times it appears as a component in $(y^1, \ldots, y^m)$, for each $a \in A$]. Let condition C' be: $[(x^1, \ldots, x^m)\ E_m^*\ (y^1, \ldots, y^m)$, $x^j \prec y^j$ or $x^j \sim y^j$ for $j = 1, \ldots, m-1] \Rightarrow$ not $x^m \prec y^m$. Show that $C' \Rightarrow C$ of Theorem 4.1 and that $C' \Rightarrow$ if $x, y \in X$ and $x_1, \ldots, x_n$ is a permutation of $y_1, \ldots, y_n$ then $(x_1, \ldots, x_n) \sim (y_1, \ldots, y_n)$.

3. With $X = A \times A$ suppose $u(a, b) = u(b, a)$ for all $a, b \in A$ and that $x \prec y \Leftrightarrow u(x) < u(y)$. With $A = \mathrm{Re}$, specify a u that satisfies these conditions (define $\prec$ from $<$) and for which there is no corresponding additive representation as in (7.1).

4. With $X = A^n$ suppose $\prec$ on A^n is a persistent weak order. Define $\prec^0$ on A by $a \prec^0 b \Leftrightarrow (x_1, \ldots, x_{i-1}, a, x_{i+1}, \ldots, x_n) \prec (x_1, \ldots, x_{i-1}, b, x_{i+1}, \ldots, x_n)$ for some $x_i \in A$. Prove

a. $\prec^0$ on A is a weak order,
b. $(x_i \prec^0 y_i$ or $x_i \sim^0 y_i$ for $i = 1, \ldots, n) \Rightarrow x \preccurlyeq y$,
c. $(x_i \preccurlyeq^0 y_i$ for all i and $x_i \prec^0 y_i$ for some $i) \Rightarrow x \prec y$.

5. Suppose $\prec$ on A^n is a strict partial order, $\prec$ is persistent, and $\prec_i$ on A is defined as in the paragraph preceding (7.3). Prove that each $\prec_i$ is a strict partial order and all $\prec_i$ are identical. Show also that when $\prec_i$ is defined in this way and $\prec$ is persistent then it is possible to have all $\prec_i$ identical weak orders on A when $\prec$ on A^n is not even a strict partial order.

6. Show that $u_1(b) - u_1(a) > u_2(b) - u_2(a) > \cdots > u_n(b) - u_n(a)$ when $\bar{a} \prec \bar{b}$, (7.1) holds, and $\prec$ is impatient.

7. Show that if $X = A^n$, $\prec^*$ on $X \times X$ is persistent, and $x \prec y \Leftrightarrow x - x \prec^* y - x$, then $\prec$ on X is persistent.

8. Show that if $X = \prod_{i=1}^n X_i$, if (7.4) holds for all $x, y, z, w \in X$ and if $x - x' \prec^* y - y' \Rightarrow z - z' \prec^* w - w'$ whenever $i \in \{1, \ldots, n\}$, $(x_j = x_j', y_j = y_j', z_j = z_j', w_j = w_j')$ for all $j \neq i$ and $(x_i, x_i', y_i, y_i') = (z_i, z_i', w_i, w_i')$ then there are real-valued u_i on X_i that satisfy $u(x) = \sum u_i(x_i)$ for all $x \in X$.

9. Show that (7.9) holds with the $\lambda_i > 0$ if and only if there are u_i satisfying (7.1) that are positive linear transformations of each other.

10. Verify the linear transformation assertion in the proof of Theorem 7.3.

11. Show that if (7.13) holds with $\pi > 0$ and if $\prec$ is impatient and $x \prec y$ for some $x, y \in A^n$, then $\pi < 1$.

12. Under the hypotheses of Theorem 7.5 does (7.14) hold for every $e \in A$?

13. Williams and Nassar (1966). Let H be the following set of hypotheses: $X = \mathrm{Re}^n$, conditions 1, 2, and 3 of Theorem 3.3, and $x \prec y \Leftrightarrow \mathbf{0} \prec y - x$, for all $x, y \in X$. The final assumption is referred to as "marginal consistency." Show that the following hold, given H.

a. $x \sim y \Leftrightarrow x - y \sim \mathbf{0}$.
b. $x \sim y \Leftrightarrow -x \sim -y$.
c. $x \sim y \Leftrightarrow x + z \sim y + z$ for every $z \in \mathrm{Re}^n$.
d. $(x \sim y, z \sim w) \Rightarrow x + z \sim y + w$.
e. $x \sim y \Rightarrow Mx \sim My$ for every integer M.

f. $x \prec y \Leftrightarrow -y \prec -x$.
g. $x \prec y \Leftrightarrow x + z \prec y + z$ for every $z \in \mathrm{Re}^n$.
h. $(x \prec y, z \preccurlyeq w) \Rightarrow x + z \prec y + w$.
i. $x \prec y \Rightarrow Mx \prec My$ for every positive integer M, and $x \prec y \Rightarrow My \prec Mx$ for every negative integer M.
j. If M is a nonzero integer then $Mx \sim My \Rightarrow x \sim y$.
k. $x \sim y \Rightarrow \alpha x \sim \alpha y$ for every rational number α.
m. $x \sim y \Rightarrow \alpha x \sim \alpha y$ for every $\alpha \in \mathrm{Re}$.

14. (*Continuation.*) Show that H implies that there are positive numbers $\lambda_1, \ldots, \lambda_n$ such that

$$x \prec y \Leftrightarrow \sum_{i=1}^{n} \lambda_i x_i < \sum_{i=1}^{n} \lambda_i y_i, \qquad \text{for all } x, y \in X. \tag{7.15}$$

To do this show first that, for each $x \in \mathrm{Re}^n$, there is one and only one $a \in \mathrm{Re}$ for which $x \sim \bar{a}$. Then take $u(x) = a$ when $x \sim \bar{a}$, so that u satisfies $x \prec y \Leftrightarrow u(x) < u(y)$. Finally, use results d and m of the preceding exercise to show that u can be written as $u(x) = \sum \lambda_i x_i$ where $\lambda_i \sim (0, \ldots, 0, 1, 0, \ldots, 0)$.

PART

II

EXPECTED-UTILITY THEORY

Until the mid twentieth century, utility theory focused on preference structures that do not explicitly incorporate uncertainty or probability, the yardstick for uncertainty. The expected-utility theory of John von Neumann and Oskar Morgenstern, and an earlier theory by Frank P. Ramsey, stimulated new interest in the role of uncertainty in preference structures.

An expected-utility theory may incorporate probabilities in the alternatives of the preference structure or it may formulate uncertainty in the alternatives without its prior encoding in terms of probability. In the latter case, probabilities as well as utilities are derived from the axioms. In the former case only utilities are derived from the axioms since the probabilities are already part of the axiomatic structure. The former approach is used in this part of the book: the alternatives are probability measures defined on a set of consequences. Basic theory is in Chapters 8, 9, and 10: additive, expected-utility theory for multiple-factor situations is in Chapter 11.

Chapter 8

EXPECTED UTILITY WITH SIMPLE PROBABILITY MEASURES

When each strategy or decision alternative corresponds to a simple probability measure on the consequences in a set X, we consider the expected-utility model for computing utilities of the strategies, or their associated measures. The idea for this model dates at least from Bernoulli (1738) but it was not until the present century that apparently reasonable preference axioms were given as a basis for the model. The axioms of this chapter are similar to those initiated by von Neumann and Morgenstern (1947) and to later modifications by Friedman and Savage (1948, 1952), Marschak (1950), Herstein and Milnor (1953), Cramer (1956), Luce and Raiffa (1957), and Blackwell and Girshick (1954). The last of these applies to probability measures that are more general than those considered in this chapter. They will be examined in Chapter 10.

After an introductory example and a brief discussion of simple probability measures we shall consider the basic theorem and then offer some criticisms of its preference conditions. A complete proof of the basic weak-order theorem is given in Section 8.4. The case of intransitive indifference is investigated in the next chapter.

8.1 EXAMPLE

Suppose that the owner of a small construction firm plans to submit a sealed bid for a job that he estimates will cost his company \$200000 to complete. If he bids \$$x$ and gets the job, he will be paid \$$x$: his profit is \$$x$ − \$200000.

Since the construction industry is in a slump, he believes that there will be many bids. From his prior experience and knowledge of the current situation he estimates the probability $p(x)$ of getting the job if he bids \$$x$. [Winkler

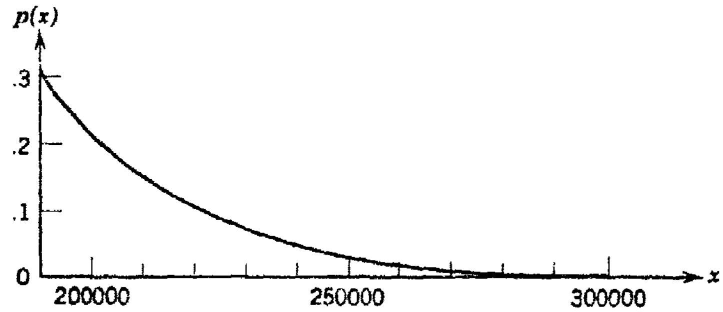

Figure 8.1 Probability of getting job for a bid of $\$x$.

(1967*a*, 1967*b*) discusses some ways of doing this.] $p(x)$ for $190000 \leq x \leq 300000$ is shown in Figure 8.1.

Because of the scarcity of work the owner would be willing to take the job at a loss of not more than \$10000. In other words, (get job and make −\$10000) $\sim$ (don't get job). Using an appropriate method of scaling utilities for the expected-utility model [see, for example, Pratt, Raiffa, and Schlaifer (1964), Swalm (1966), or Fishburn (1967)], the owner estimates his utility function for net profit (assuming he gets the job) as shown in Figure 8.2. The figure indicates that he is indifferent between making \$10000 with certainty and a 50-50 gamble giving either −\$10000 or \$100000. He is indifferent also between making \$50000 with certainty and an 80-20 gamble giving \$100000 (with probability .8) or −\$10000 (with probability .2). According to the

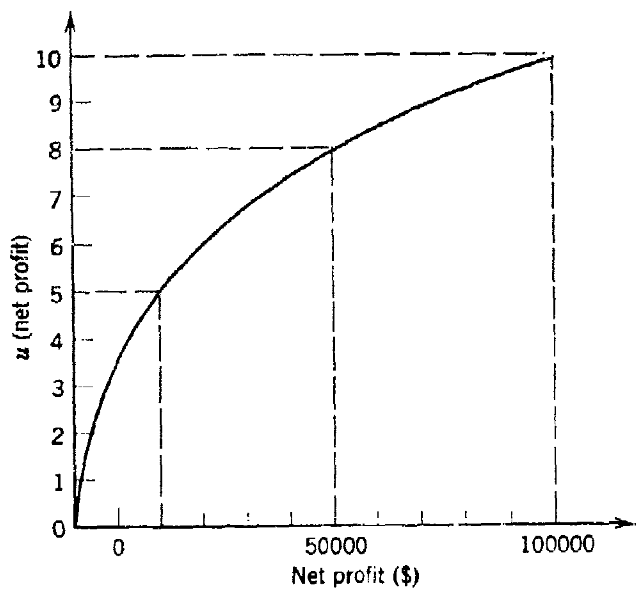

Figure 8.2 Utility of net profit.

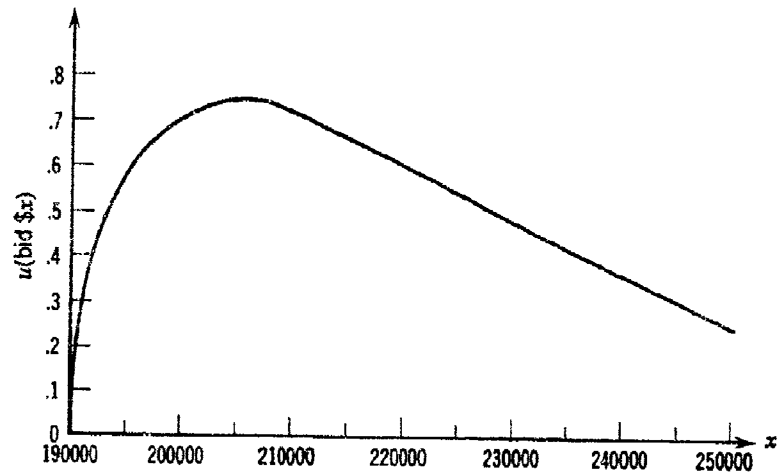

Figure 8.3 Expected utilities using Figures 8.1 and 8.2.

expected-utility model, the latter indifference comparison transforms into $u(\$50000) = .8u(\$100000) + .2u(-\$10000)$. Equations such as this can be used as a guide in constructing and checking u.

If he bids $\$x$ his expected utility will be $p(x)$ u(get the job and make $\$x - \$200000$ net profit) + $[1 - p(x)]u$(don't get job). By Figure 8.2 and (get job and make $-\$10000$) $\sim$ (don't get job), u(don't get job) $= 0$ so that

$$u(\text{bid } \$x) = p(x)u(\text{get job and make } \$x - \$200000 \text{ net profit}).$$

Reading off approximate values for $p(x)$ and $u(\$x - \$200000)$ from Figures 8.1 and 8.2 we obtain the expected-utility curve in Figure 8.3, which shows that expected utility is maximized at about $x = 206000$. A bid of about $206000 is therefore recommended.

8.2 SIMPLE PROBABILITY MEASURES

Definition 8.1. A simple probability measure on X is a real-valued function P defined on the set of all subsets of X such that

1. $P(A) \geq 0$ for every $A \subseteq X$,
2. $P(X) = 1$,
3. $P(A \cup B) = P(A) + P(B)$ when $A, B \subseteq X$ and $A \cap B = \varnothing$,
4. $P(A) = 1$ for some *finite* $A \subseteq X$.

Property (4) distinguishes P as a *simple* probability measure. Chapter 10 removes this restriction and considers expected utility for more general measures.

Property (3) is the finite additivity property: the probability of the union of two disjoint subsets of X equals the sum of the two separate probabilities.

$P(\{x\})$, which we shall write as $P(x)$, is the probability assigned by P to the unit subset $\{x\}$ of X.

THEOREM 8.1. *Suppose P is a simple probability measure on X. Then $P(x) = 0$ for all but a finite number of $x \in X$ and, for all $A \subseteq X$,*

$$P(A) = \sum_{x \in A} P(x). \tag{8.1}$$

Proof. Suppose P is simple and A is a finite subset of X for which $P(A) = 1$. Then $P(x) = 0$ for all $x \notin A$, for otherwise, if $P(x) > 0$, $P(A \cup \{x\}) > 1$ by (3) of Definition 8.1, which by (1) and (3) then leads to $P(X) > 1$, contradicting (2). By successive uses of (3), (8.1) holds when A is finite. For arbitrary $A \subseteq X$ let $B = \{x : x \in A, P(x) > 0\}$ and $C = \{x : x \in A, P(x) = 0\}$. By (3), $P(A) = P(B) + P(C)$. Moreover, B is finite so that (8.1) holds if $P(C) = 0$. If $P(C) > 0$ then, by (3), $P(C \cup \{x : x \in X, P(x) > 0\}) > 1$ since if $P\{x : x \in X, P(x) > 0\} < 1$ then, by (8.1) for finite sets, $P(D) < 1$ for every finite $D \subseteq X$. Hence, if $P(C) > 0$ we find again that $P(X) > 1$. ◆

Convex Combinations of Measures

In expected-utility theory we use a rule for combining two probability measures to form a third measure. This rule can of course be extended to the combination of any finite number of measures.

Definition 8.2. If P and Q are simple probability measures on X and $\alpha \in [0, 1]$ then $\alpha P + (1 - \alpha)Q$ is the function that assigns the number $\alpha P(A) + (1 - \alpha)Q(A)$ to each $A \subseteq X$.

Under the definition's hypotheses it is readily seen that $\alpha P + (1 - \alpha)Q$ is a simple probability measure on X.

If $P(\$100) = .3$, $P(\$200) = .7$, $Q(\$100) = .5$, and $Q(\$300) = .5$ then, with $R = .1P + .9Q$, $R(\$100) = .48$, $R(\$200) = .07$, and $R(\$300) = .45$.

Expected Value

If P is a simple probability measure on X and f is a real-valued function on X then the so-called expected value of f with respect to P, written here as $E(f, P)$, is defined by

$$E(f, P) = \sum_{x \in X} f(x)P(x). \tag{8.2}$$

With P, Q, and R as in the preceding paragraph and with $f(x) = x$, $E(f, P) = \$170$, $E(f, Q) = \$200$, and $E(f, R) = \$197 = .1E(f, P) + .9E(f, Q)$. In general, $E(f, \alpha P + (1 - \alpha)Q) = \alpha E(f, P) + (1 - \alpha)E(f, Q)$.

8.3 EXPECTED UTILITY FOR SIMPLE MEASURES

If $\mathcal{P}_s$ is the set of all simple probability measures on X then the measures that correspond to the strategies in the type of situation considered in this chapter comprise a subset of $\mathcal{P}_s$. In our preference conditions for expected utility we shall use all distributions in $\mathcal{P}_s$ for two related reasons. The first is for mathematical expediency, for when $\mathcal{P}_s$ is used it is closed under convex combinations as defined by Definition 8.3: if $P, Q \in \mathcal{P}_s$ and $\alpha \in [0, 1]$ then $\alpha P + (1 - \alpha)Q \in \mathcal{P}_s$. The second reason concerns the estimation of utilities, for when the theory is used as a basis for estimating u on X it is often convenient to use measures in $\mathcal{P}_s$ that have $P(x) > 0$ for only one to two $x \in X$, and such measures may correspond to no actual strategies.

The following theorem will be seen to be a corollary of a more general theorem that is presented and proved in the next section.

THEOREM 8.2. *Suppose that $\mathcal{P}_s$ is the set of all simple probability measures on X and $\prec$ is a binary relation on $\mathcal{P}_s$. Then there is a real-valued function u on X that satisfies*

$$P \prec Q \Leftrightarrow E(u, P) < E(u, Q), \qquad \text{for all } P, Q \in \mathcal{P}_s \tag{8.3}$$

if and only if, for all $P, Q, R \in \mathcal{P}_s$,

1. *$\prec$ on $\mathcal{P}_s$ is a weak order,*
2. $(P \prec Q, 0 < \alpha < 1) \Rightarrow \alpha P + (1 - \alpha)R \prec \alpha Q + (1 - \alpha)R,$
3. $(P \prec Q, Q \prec R) \Rightarrow \alpha P + (1 - \alpha)R \prec Q$ *and* $Q \prec \beta P + (1 - \beta)R$ *for some* $\alpha, \beta \in (0, 1)$.

Moreover, u in (8.3) is unique up to a positive linear transformation: that is, if u satisfies (8.3) then a real-valued function v on X satisfies $P \prec Q \Leftrightarrow E(v, P) < E(v, Q)$, for all $P, Q \in \mathcal{P}_s$, if and only if there are numbers $a > 0$ and b such that

$$v(x) = au(x) + b \qquad \text{for all } x \in X. \tag{8.4}$$

Suppose we extend u to $\mathcal{P}_s$ by defining $u(P) = E(u, P)$. Then, if (8.3) holds, $P \prec Q \Leftrightarrow u(P) < u(Q)$. Now if v on $\mathcal{P}_s$ is *any* order-preserving (not necessarily linear) transformation of u on $\mathcal{P}_s$, then $P \prec Q \Leftrightarrow v(P) < v(Q)$. Given such a v we can define v on X by $v(x) = v(P)$ when $P(x) = 1$. However, if v is not a *linear* transformation of u then $v(P) = E(v, P)$ must be false for some $P \in \mathcal{P}_s$. In other words there are functions v on $\mathcal{P}_s$ that satisfy $P \prec Q \Leftrightarrow v(P) < v(Q)$ but do not satisfy $P \prec Q \Leftrightarrow E(v, P) < E(v, Q)$ when v on X is defined from v on $\mathcal{P}_s$ in the manner indicated (provided that $P \prec Q$ for some $P, Q \in \mathcal{P}_s$).

Condition 1: Weak Order

Condition 1, weak order, can easily be criticized for its implication of transitive indifference. For example, let consequences be amounts of money viewed as potential increments to one's present wealth. Let $P(\$35) = 1$, $Q(\$36) = 1$, and $R(\$0) = R(\$100) = .5$. Surely $P < Q$. But it seems quite possible that $P \sim R$ and $Q \sim R$, in which case $\sim$ is not transitive.

For this reason the next chapter examines the case where $<$ on $\mathcal{P}_s$ is only assumed to be a strict partial order. We shall not consider interval orders and semiorders *per se*, as in Chapter 2, for conditions $p10$ and $p11$ of Section 2.4 are liable to criticisms of the sort given above. For example, if $Q'(\$35.50) = 1$, then $P < Q' < Q$ but R might be indifferent to each of these, which would violate $p11$. Moreover, if $<$ on $\mathcal{P}_s$ is assumed to be irreflexive and to satisfy $p11$, and if condition 2 of Theorem 8.2 holds then $\sim$ on $\mathcal{P}_s$ is transitive. For suppose to the contrary that $(P \sim Q, Q \sim R, P < R)$. Then, by condition 2 on $P < R$, $P = \frac{1}{2}P + \frac{1}{2}P < \frac{1}{2}P + \frac{1}{2}R$ and $\frac{1}{2}P + \frac{1}{2}R < \frac{1}{2}R + \frac{1}{2}R = R$, so that, by $p11$, $P < Q$ or $Q < R$ which contradicts $(P \sim Q, Q \sim R)$.

Condition 2: Independence

Condition 2, a form of independence axiom, is regarded by many as the core of expected-utility theory, for without it the "expectation" part of expected utility vanishes. Moreover, this condition is often regarded as a principal normative criterion of the theory, along with transitivity of $<$.

$\alpha P + (1 - \alpha)R$ may be viewed in two ways: either as a gamble that yields $x \in X$ with probability $\alpha P(x) + (1 - \alpha)R(x)$, or as a two-stage process whereby P (or R) is selected in the first stage with probability α (or $1 - \alpha$) and then x is selected at the second stage using the one of P and R already selected. These two interpretations are probabilistically identical although they are not psychologically identical. For example, you might find the two-stage process more exciting.

As a normative criterion, $(P < Q, 0 < \alpha < 1) \Rightarrow \alpha P + (1 - \alpha)R < \alpha Q + (1 - \alpha)R$ is usually defended with the two-stage argument. If you prefer Q to P then it seems reasonable in view of the two-stage interpretation that you should prefer $\alpha Q + (1 - \alpha)R$ to $\alpha P + (1 - \alpha)R$, or that, in the following "payoff" matrix, you should prefer A to B when you have a choice between

	α	$1 - \alpha$
Option A	Q	R
Option B	P	R

A and B and, independent of your choice, a "coin" with probability α for "heads" and probability $1 - \alpha$ for "tails" is flipped to determine the appropriate column.

Condition 2 has several related functions as a guide in making consistent preference judgments. First, it may help to uncover preferences between more complex alternatives on the basis of preferences between simpler alternatives. Suppose that, initially, a person has no clear preference between R and S where

$$R(\$50) = .10, \quad R(\$80) = .45, \quad R(\$100) = .45$$
$$S(\$0) = .02, \quad S(\$80) = .45, \quad S(\$100) = .53,$$

but definitely prefers Q to P when $Q(\$0) = .2$, $Q(\$100) = .8$, and $P(\$50) = 1$. Let $T(\$80) = T(\$100) = .5$. In view of the fact that $S = .1Q + .9T$ and $R = .1P + .9T$, his preference for Q over P may convince him that he should prefer S to R even though he might feel that S and R are "very close together."

Condition 2 can also be useful in uncovering inconsistencies in preference judgments. Consider an example used by Savage (1954, pp. 101–103) that is due to Allais (1953). Which of Q and P do you prefer?

$$Q(\$500000) = 1; \quad P(\$2500000) = .10,$$
$$P(\$500000) = .89, \quad P(\$0) = .01.$$

Also, which of R and S do you prefer?

$$R(\$500000) = .11, \quad R(\$0) = .89; \quad S(\$2500000) = .10, \quad S(\$0) = .90$$

According to Allais and Savage it is not unusual to find $P < Q$ and $R < S$. Now with $T(\$2500000) = \frac{10}{11}$, $T(\$0) = \frac{1}{11}$, and $V(\$0) = 1$,

$$Q = .11Q + .89Q$$
$$P = .11T + .89Q$$

and

$$R = .11Q + .89V$$
$$S = .11T + .89V.$$

Since condition 2 implies the converse of itself in the presence of the other conditions, $P < Q \Rightarrow T < Q$ and $R < S \Rightarrow Q < T$, so that an "inconsistency" has been uncovered. In Allais' viewpoint, this result speaks against the reasonableness of condition 2. On the other hand, Savage suggests that many people would be alarmed at the apparent inconsistency and, accepting the "reasonableness" of condition 2, wish to revise their initial judgments so that the revisions are consistent with the condition.

Condition 3: An Archimedean Axiom

The third condition in Theorem 8.2 says that if $P < Q < R$ then there is some nontrivial mixture of P and R that is less preferred than Q, and also

some nontrivial mixture of P and R that is preferred to Q. It specifically prohibits the possibility that not $\alpha P + (1 - \alpha)R < Q$ for all $\alpha \in (0, 1)$, or that not $Q < \alpha P + (1 - \alpha)R$ for all $\alpha \in (0, 1)$ when $P < Q < R$.

Suppose that a newly minted penny will be flipped n times and that, for any positive α, you feel that there is an $n(\alpha)$ such that α exceeds the probability that every one of the $n(\alpha)$ flips will result in a head. Consider a choice between A and B:

A. Receive \$1 regardless of the results of the n flips,
B. Be executed if every flip results in a head, and receive \$2 otherwise.

If execution $<$ \$1 $<$ \$2 and if you prefer A to B regardless of how large n is taken to be, then you violate condition 3. If the coin is flipped 100 times, then under B there is only one sequence of the more than 1,000,000,000,000,000,000,000,000,000,000 possible sequences under which you would be executed. In view of such numbers, many people might find a satisfactorily large value of n for which they would choose B. It is often claimed that the willingness that many people show toward small risks such as crossing the street or driving a car is sufficiently convincing evidence in favor of the condition.

Despite the fact that condition 3 is called an Archimedean axiom, it and weak order do not imply the existence of a u on $\mathcal{P}_s$ that satisfies $P < Q \Leftrightarrow u(P) < u(Q)$. In other words, conditions 1 and 3 do not imply (see Theorem 3.1) that $\mathcal{P}_s/\sim$ includes a countable subset that is order dense in $\mathcal{P}_s/\sim$. Exercise 6 goes into this further.

Hausner (1954) considers the case where condition 3 is not assumed to hold. To conditions 1 and 2 he adds the indifference version of condition 2, $(P \sim Q, 0 < \alpha < 1) \Rightarrow \alpha P + (1 - \alpha)R \sim \alpha Q + (1 - \alpha)R$, which as we shall see in the next section is implied by conditions 1, 2, and 3. His axioms imply a lexicographic form of expected utility, but the dimensionality of this form might not be finite. In the 2-dimensional case his representation would be $P < Q \Leftrightarrow (E(u_1, P), E(u_2, P)) <^L (E(u_1, Q), E(u_2, Q))$ where u_1 and u_2 are real-valued functions on X and $<^L$ is defined as in (4.10).

8.4 MIXTURE SETS

We shall now develop and prove a theorem that is more general than Theorem 8.2. The reason for this is that the more general theorem will be used in later developments, especially in Chapter 13. The generalization uses Herstein and Milnor's (1953) definition of a mixture set.

Definition 8.3. A *mixture set* is a set $\mathcal{P}$ and a function that assigns an element $\alpha P + (1 - \alpha)Q$ in $\mathcal{P}$ to each $\alpha \in [0, 1]$ and each $(P, Q) \in \mathcal{P} \times \mathcal{P}$ such

that, for all $P, Q \in \mathcal{P}$ and $\alpha, \beta \in [0, 1]$,

$M1.$ $1P + 0Q = P$,
$M2.$ $\alpha P + (1-\alpha)Q = (1-\alpha)Q + \alpha P$,
$M3.$ $\alpha[\beta P + (1-\beta)Q] + (1-\alpha)Q = \alpha\beta P + (1-\alpha\beta)Q$.

The $\mathcal{P}_s$ with $\alpha P + (1-\alpha)Q$ as in Definition 8.2 is a mixture set. Along with $M1$ through $M3$ we shall use the following:

$M4.$ $\alpha P + (1-\alpha)P = P$,
$M5.$ $\alpha[\beta Q + (1-\beta)R] + (1-\alpha)[\gamma Q + (1-\gamma)R]$

$$= [\alpha\beta + (1-\alpha)\gamma]Q + [\alpha(1-\beta) + (1-\alpha)(1-\gamma)]R.$$

The first of these follows from $M1$–$M3$ as follows: $\alpha P + (1-\alpha)P = \alpha[1P + 0P] + (1-\alpha)P = \alpha[0P + 1P] + (1-\alpha)P = 0P + 1P = 1P + 0P = P$. The second follows easily from $M1$–$M3$ if β or γ equals 0 or 1. Henceforth, to verify $M5$ for a mixture set, we suppose that $\beta, \gamma \in (0, 1)$ and that $\beta \leq \gamma$ for definiteness. Following Luce and Suppes (1965, p. 288):

$$\begin{aligned}
&[\alpha\beta + (1-\alpha)\gamma]Q + [\alpha(1-\beta) + (1-\alpha)(1-\gamma)]R \\
&= \{[\alpha\beta/\gamma + (1-\alpha)]\gamma\}Q + \{1 - [\alpha\beta/\gamma + (1-\alpha)]\gamma\}R \\
&= [\alpha\beta/\gamma + (1-\alpha)][\gamma Q + (1-\gamma)R] + [1 - \alpha\beta/\gamma - (1-\alpha)]R \\
&\qquad\qquad \text{by } M3 \\
&= [\alpha(1-\beta/\gamma)]R + [1 - \alpha(1-\beta/\gamma)][\gamma Q + (1-\gamma)R] \qquad \text{by } M2 \\
&= \alpha\{(1-\beta/\gamma)R + (\beta/\gamma)[\gamma Q + (1-\gamma)R]\} + (1-\alpha)[\gamma Q + (1-\gamma)R] \\
&\qquad\qquad \text{by } M3 \\
&= \alpha\{(\beta/\gamma)[\gamma Q + (1-\gamma)R] + (1-\beta/\gamma)R\} + (1-\alpha)[\gamma Q + (1-\gamma)R] \\
&\qquad\qquad \text{by } M2 \\
&= \alpha[\beta Q + (1-\beta)R] + (1-\alpha)[\gamma Q + (1-\gamma)R] \qquad \text{by } M3.
\end{aligned}$$

As a preface to the main theorem we consider a succession of lemmas, as incorporated in the following theorem. Conclusion 5 of the theorem is due to Jensen (1967). $\sim$ and $\precsim$ are defined as in (2.2) and (2.3).

THEOREM 8.3. *Suppose that $\mathcal{P}$ is a mixture set and that the following hold for all $P, Q, R \in \mathcal{P}$:*

A1. $\prec$ *on $\mathcal{P}$ is a weak order,*
A2. $(P \prec Q, 0 < \alpha < 1) \Rightarrow \alpha P + (1-\alpha)R \prec \alpha Q + (1-\alpha)R$,
A3. $(P \prec Q,\ Q \prec R) \Rightarrow \alpha P + (1-\alpha)R \prec Q$ *and* $Q \prec \beta P + (1-\beta)R$ *for some $\alpha, \beta \in (0, 1)$. Then, for all $P, Q, R, S \in \mathcal{P}$,*

$C1$. $(P < Q, 0 \le \alpha < \beta \le 1) \Rightarrow \beta P + (1-\beta)Q < \alpha P + (1-\alpha)Q$,

$C2$. $(P \preccurlyeq Q,\ Q \preccurlyeq R,\ P < R) \Rightarrow Q \sim \alpha P + (1-\alpha)R$ *for exactly one* $\alpha \in [0, 1]$,

$C3$. $(P < Q, R < S, 0 \le \alpha \le 1) \Rightarrow \alpha P + (1-\alpha)R < \alpha Q + (1-\alpha)S$,

$C4$. $(P \sim Q, 0 \le \alpha \le 1) \Rightarrow \alpha P + (1-\alpha)Q \sim P$,

$C5$. $(P \sim Q, 0 \le \alpha \le 1) \Rightarrow \alpha P + (1-\alpha)R \sim \alpha Q + (1-\alpha)R$.

Proofs. $C1$. If $\beta < 1$, $\beta P + (1-\beta)Q < \beta Q + (1-\beta)Q$ by $A2$, and hence $\beta P + (1-\beta)Q < Q$ by $M4$. If $\beta = 1$, then $\beta P + (1-\beta)Q < Q$ by $M1$. If $0 < \alpha$, then $(\alpha/\beta)[\beta P + (1-\beta)Q] + (1-\alpha/\beta)[\beta P + (1-\beta)Q] < (\alpha/\beta)[\beta P + (1-\beta)Q] + (1-\alpha/\beta)Q$ by $A2$, and hence $\beta P + (1-\beta)Q < \alpha P + (1-\alpha)Q$ by $M3$ and $M4$. If $\alpha = 0$, $\beta P + (1-\beta)Q < \alpha P + (1-\alpha)Q$ by $M1$ and $M2$.

$C2$. Suppose first that $Q \sim P$. Then $Q \sim 1P + 0R$ by $M1$, and $1P + 0R < \beta P + (1-\beta)R$ for every $\beta < 1$ by $C1$ and $M2$. Then, by transitivity (see Theorem 2.1*d*), $\alpha = 1$ is the unique $\alpha \in [0, 1]$ for which $Q \sim \alpha P + (1-\alpha)R$. A symmetric proof holds if $Q \sim R$ (in which case $\alpha = 0$). Finally, if $P < Q < R$, the proof of Lemma 3.1 applies with the obvious notational changes and the use of $C1$, $A1$–$A3$, and $M2$ and $M3$.

$C3$. If $0 < \alpha < 1$, $\alpha P + (1-\alpha)R < \alpha Q + (1-\alpha)R$ and $(1-\alpha)R + \alpha Q < (1-\alpha)S + \alpha Q$ by $A2$.

$C4$. Suppose $P \sim Q$ and $\alpha P + (1-\alpha)Q < P$. Then $\alpha P + (1-\alpha)Q < Q$ (Theorem 2.1*d*). Then, by $C3$, $\alpha[\alpha P + (1-\alpha)Q] + (1-\alpha)[\alpha P + (1-\alpha)Q] < \alpha P + (1-\alpha)Q$ or, by $M4$, $\alpha P + (1-\alpha)Q < \alpha P + (1-\alpha)Q$, which is false. Similarly, not $(P \sim Q, P < \alpha P + (1-\alpha)Q)$. Hence $P \sim Q \Rightarrow \alpha P + (1-\alpha)Q \sim P$.

$C5$. $M1$ and $M2$ yield the conclusion if $\alpha \in \{0, 1\}$. Take $(P \sim Q, 0 < \alpha < 1)$. If $R \sim P$ then, by $C4$, $\alpha P + (1-\alpha)R \sim P \sim Q \sim \alpha Q + (1-\alpha)R$, or $\alpha P + (1-\alpha)R \sim \alpha Q + (1-\alpha)R$. Henceforth take $R < P$ (the $P < R$ proof is similar). Then $R < \alpha P + (1-\alpha)R$ by $C1$ and $M4$. Suppose also that $\alpha P + (1-\alpha)R < \alpha Q + (1-\alpha)R$. Then, by $C2$, $\alpha P + (1-\alpha)R \sim (1-\beta)R + \beta[\alpha Q + (1-\alpha)R]$ for a unique $\beta \in (0, 1)$. Hence $\alpha P + (1-\alpha)R \sim \alpha\beta Q + (1-\alpha\beta)R$ by $M2$ and $M3$. Also, since $R < Q$, $(1-\beta)R + \beta Q < Q \sim P$ by $C1$ and $M4$: hence $\beta Q + (1-\beta)R < P$ by $A1$ and $M2$: then $\alpha[\beta Q + (1-\beta)R] + (1-\alpha)R < \alpha P + (1-\alpha)R$ by $A2$: finally, $\alpha\beta Q + (1-\alpha\beta)R < \alpha P + (1-\alpha)R$ by $M3$, thus contradicting $\alpha P + (1-\alpha)R \sim \alpha\beta Q + (1-\alpha\beta)R$. Hence $\alpha P + (1-\alpha)R < \alpha Q + (1-\alpha)R$ is false. Similarly $\alpha Q + (1-\alpha)R < \alpha P + (1-\alpha)R$ is false. Hence $\alpha P + (1-\alpha)R \sim \alpha Q + (1-\alpha)R$. ◆

The Main Theorem

THEOREM 8.4. *Suppose $\mathscr{S}$ is a mixture set. Then $A1$, $A2$, and $A3$ of Theorem 8.3 hold for all P, Q, $R \in \mathscr{S}$ if and only if there is a real-valued function u on*

$\mathcal{P}$ *such that*

$$P < Q \Leftrightarrow u(P) < u(Q), \quad \text{for all } P, Q \in \mathcal{P} \tag{8.5}$$

$$u(\alpha P + (1-\alpha)Q) = \alpha u(P) + (1-\alpha)u(Q), \quad \text{for all } (\alpha, P, Q) \in [0, 1] \times \mathcal{P}^2. \tag{8.6}$$

Moreover, if u on $\mathcal{P}$ satisfies (8.5) *and* (8.6) *then a real-valued function v on $\mathcal{P}$ satisfies* (8.5) *and* (8.6) *with u replaced by v if and only if there are numbers $a > 0$ and b such that*

$$v(P) = au(P) + b \quad \text{for all } P \in \mathcal{P}. \tag{8.7}$$

Theorem 8.2 results from this when $\mathcal{P} = \mathcal{P}_s$ and u on X is defined from u on $\mathcal{P}$ by $u(x) = u(P)$ when $P(x) = 1$. If $\{x:P(x) > 0\} = \{x_1, \ldots, x_n\}$ then repeated applications of (8.6) with $P_i \in \mathcal{P}_s$ such that $P_i(x_i) = 1$ give $u(P) = u(\sum_{i=1}^n P(x_i)P_i) = \sum_{i=1}^n P(x_i)u(x_i) = E(u, P)$, so that $P < Q \Leftrightarrow E(u, P) < E(u, Q)$ by (8.5). (8.4) follows from (8.7).

The necessity of $A1$, $A2$, and $A3$ for (8.5) and (8.6) is obvious. To prove sufficiency, Part I of the following proof shows that (8.5) and (8.6) hold on $RS = \{P: R \leqslant P \leqslant S\}$ when $R < S$. We assume $R < S$ for some $R, S \in \mathcal{P}$ for otherwise the conclusion is obvious. Part II extends (8.5) and (8.6) to all of $\mathcal{P}$. Part III verifies (8.7).

Proof, Part I. Assume that $A1$, $A2$, and $A3$ hold and that $R < S$. Let $RS = \{P: P \in \mathcal{P}, R \leqslant P \leqslant S\}$. By $C2$ there is a unique number $f(P) \in [0, 1]$ for each $P \in RS$ such that

$$P \sim [1 - f(P)]R + f(P)S, \quad \text{with} \quad f(R) = 0 \quad \text{and} \quad f(S) = 1. \tag{8.8}$$

Suppose $P, Q \in RS$ and $f(P) < f(Q)$. Then, by $C1$, $[1 - f(P)]R + f(P)S < [1 - f(Q)]R + f(Q)S$. Transitivity and (8.8) then give $P < Q$. On the other hand, if $f(P) = f(Q)$ then (8.8) and transitivity imply $P \sim Q$. Thus

$$P < Q \Leftrightarrow f(P) < f(Q), \quad \text{for all } P, Q \in RS. \tag{8.9}$$

If $P, Q \in RS$ and $\alpha \in [0, 1]$ then $\alpha P + (1-\alpha)Q \in RS$. If $\alpha \in \{0, 1\}$ this follows from $M1$ and $M2$. If $0 < \alpha < 1$ then $R = \alpha R + (1-\alpha)R \leqslant \alpha P + (1-\alpha)R = (1-\alpha)R + \alpha P \leqslant (1-\alpha)Q + \alpha P = \alpha P + (1-\alpha)Q \leqslant \alpha S + (1-\alpha)Q = (1-\alpha)Q + \alpha S \leqslant (1-\alpha)S + \alpha S = S$ by $M4$, $A2$ or $C5$, $M2$, $A2$ or $C5$, $M2$, $A2$ or $C5$, and $M4$, in that order.

Therefore, if $P, Q \in RS$ and $\alpha \in [0, 1]$ then, by (8.8),

$$\alpha P + (1-\alpha)Q \sim [1 - f(\alpha P + (1-\alpha)Q)]R + f(\alpha P + (1-\alpha)Q)S. \tag{8.10}$$

In addition, by two applications of $C5$,

$$\alpha P + (1-\alpha)Q \sim \alpha\{[1-f(P)]R + f(P)S\} \\ + (1-\alpha)\{[1-f(Q)]R + f(Q)S\},$$

so that, by $M5$,

$$\alpha P + (1-\alpha)Q \sim [1 - \alpha f(P) - (1-\alpha)f(Q)]R \\ + [\alpha f(P) + (1-\alpha)f(Q)]S.$$

From this, (8.10), transitivity, and $C1$ it follows that

$$f(\alpha P + (1-\alpha)Q) = \alpha f(P) + (1-\alpha)f(Q), \quad \text{for all } (\alpha, P, Q) \in [0,1] \times RS^2. \tag{8.11}$$

(8.9) and (8.11) verify (8.5) and (8.6) on RS.

Proof, Part II. To extend this to all of $\mathcal{T}$, fix RS with $R < S$, and let $R_iS_i = \{P : P \in \mathcal{T}, R_i \leqslant P \leqslant S_i\}$ be such that $RS \subseteq R_iS_i$ for $i = 1, 2$. Let f_i^* on R_iS_i satisfy (8.5) and (8.6) for $(\alpha, P, Q) \in [0, 1] \times R_iS_i^2$, as guaranteed by Part I. Let f_i be a positive linear transformation of f_i^* so that $f_i(R) = 0$ and $f_i(S) = 1$ for $i = 1, 2$. The f_i must satisfy (8.5) and (8.6) for $(\alpha, P, Q) \in [0, 1] \times R_iS_i^2$.

Suppose $P \in R_1S_1 \cap R_2S_2$. If $P \sim R$ or $P \sim S$ then $f_1(P) = f_2(P)$ by the definitions. Three possibilities remain as shown here with the unique element in (0, 1) as guaranteed by $C2$ and strict preference:

$$P < R < S, \qquad R \sim (1-\alpha)P + \alpha S \tag{8.12}$$

$$R < P < S, \qquad P \sim (1-\beta)R + \beta S \tag{8.13}$$

$$R < S < P, \qquad S \sim (1-\gamma)R + \gamma P. \tag{8.14}$$

Using (8.5) and (8.6) on each of these we get, for $i = 1, 2$,

$$0 = (1-\alpha)f_i(P) + \alpha \qquad (\alpha \neq 1) \tag{8.12*}$$

$$f_i(P) = \beta \tag{8.13*}$$

$$1 = \gamma f_i(P) \qquad (\gamma \neq 0) \tag{8.14*}$$

respectively, so that $f_1(P) = f_2(P)$ in each case.

Finally, let $u(P)$ be the common value of $f_i(P)$, as assured by the foregoing, for every interval of the form R_iS_i containing P, R, and S. Since every pair $P, Q \in \mathcal{T}$ is in at least one such interval it follows that u is defined on all of $\mathcal{T}$ and satisfies (8.5) and (8.6).

Proof, Part III. If u satisfies (8.5) and (8.6) and v satisfies (8.7) with $a > 0$ then v obviously satisfies (8.5) and (8.6). To go the other way, suppose v satisfies (8.5) and (8.6) along with u. If u is constant on $\mathcal{T}$ then so is v and they

are related by the positive linear transformation $v(P) = u(P) + (c' - c)$ where $u \equiv c$, $v \equiv c'$. On the other hand suppose that $R < S$ for some $R, S \in \mathcal{T}$. With such R and S fixed let

$$f_1(P) = \frac{u(P) - u(R)}{u(S) - u(R)}, \qquad f_2(P) = \frac{v(P) - v(R)}{v(S) - v(R)} \qquad \text{for all } P \in \mathcal{T}. \quad (8.15)$$

Since f_1 and f_2 are positive linear transformations of u and v, both satisfy (8.5) and (8.6). Moreover $f_1(R) = f_2(R) = 0$ and $f_1(S) = f_2(S) = 1$. If $P \sim R$ or $P \sim S$ then $f_1(P) = f_2(P)$. Or if (8.1k) holds then $f_1(P) = f_2(P)$ by (8.1k^*) for $k = 2, 3, 4$. Hence $f_1 \equiv f_2$. Then, by (8.15),

$$v(P) = \frac{v(S) - v(R)}{u(S) - u(R)} u(P) + v(R) - u(R) \frac{v(S) - v(R)}{u(S) - u(R)}$$

so that v is a positive linear transformation of u. ◆

8.5 SUMMARY

When a decision alternative has positive probability of resulting in any consequence in a finite subset of consequences and the probabilities sum to one, then a simple probability measure on X corresponds to the alternative. Three preference conditions—weak order, independence, Archimedean—for $<$ on the set of simple probability measures imply that the utility of any measure can be computed as the expected utility of the consequences with respect to that measure, provided that the consequence utilities are defined in a manner consistent with the expected-utility model.

For a general theory we defined the notion of a mixture set and applied the three conditions to it. The expected-utility model for simple probability measures illustrates one application of the general theory. Other uses of the general theory occur later.

INDEX TO EXERCISES

1. Expected net profit. 2. Simple measures. 3. Unbounded utility. 4. Positive linear transformations. 5. Independence condition. 6. Order denseness. 7. Independence. 8. Necessary conditions. 9. Expected utility. 10–11. Sequential analysis. 12–13. Certainty equivalents. 14. Pfanzagl's "consistency" axiom. 15. Linear additivity. 16. Buying and selling prices.

Exercises

1. Using Figure 8.1, sketch a curve of the expected net profit of x, similar to Figure 8.3. Approximately what x value maximizes expected net profit? Why does this differ from the x that maximizes expected utility?

2. Use (3) of Definition 8.1 to show that (*a*) $P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$ if $A_i \cap A_j = \varnothing$ whenever $i \neq j$; (*b*) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

3. Show that (8.3) does not imply that u is bounded.

4. Let u on $X = \{x, y, z, w\}$ satisfy (8.3) with $(u(x), u(y), u(z), u(w)) = (0, 1, 2, 5)$. Assuming that v satisfies (8.4) compute v on X when (*a*) $v(x) = -1$, $v(y) = 1$; (*b*) $v(x) = -10$, $v(z) = 50$; (*c*) $v(w) = 2$ and $v(x) + v(y) + v(z) + v(w) = 1$; (*d*) $v(x)v(w) = v(y)v(z) = 150$.

5. Consider P and Q as defined on \$ by the probability matrix:

	\$10	\$30	\$50	\$100	\$150
P	.2	.3	.2	.1	.2
Q	.4	.1	.1	.3	.1

Consider also two gambles for a four-ticket lottery as described in the following payoff matrix:

	Number on drawn ticket is 1 or 2	3	4
Gamble A	\$30	\$50	\$150
Gamble B	\$10	\$100	\$100

If each ticket has the same chance of being drawn, show that condition 2 of Theorem 8.2 implies $P \prec Q$ if $A \prec B$, and $Q \prec P$ if $B \prec A$. (Compute α and R that satisfy $P = \alpha A' + (1 - \alpha)R$ and $Q = \alpha B' + (1 - \alpha)R$, where A' and B' are the measures for A and B.)

6. With $\prec'$ defined on $\mathcal{P}_s/\sim$ as in (2.4) let condition 4 be: *there is a countable subset of* $\mathcal{P}_s/\sim$ *that is* $\prec'$*-order dense in* $\mathcal{P}_s/\sim$.
   a. Show that condition 1 (of Theorem 8.2) and condition 4 do not imply condition 3. (Define $\prec$ by $P \prec Q \prec R \sim S$ where $P(x) = Q(y) = 1$ for $x, y \in X$ and R and S are any two measures in $\mathcal{P}_s - \{P, Q\}$.)
   b. Show that conditions 1 and 3 do not imply condition 4. (Let $X = \{x, y\}$, let $\mathcal{P}_s$ be represented by $[0, 1]$ where $p \in [0, 1]$ is the probability assigned to x, and let $A = \{p : 0 \leq p \leq 1/2, p \text{ is rational}\}$, $B = \{p : 0 < p < 1, p \text{ is irrational}\}$, $C = \{p : 1/2 < p \leq 1, p \text{ is rational}\}$. Define $\prec$ by: $p \sim q$ if $p, q \in A$ or $p, q \in C$ or $p = q$; $p \prec q$ if ($p \in A, q \notin A$) or ($p \notin C, q \in C$) or ($p, q \in B$ and $|p - 1/2| < |q - 1/2|$) or ($p, q \in B, p < q$, and $|p - 1/2| = |q - 1/2|$).)
   c. Show that conditions 1, 3, and 4 do not imply condition 2. (Define $\prec$ by $P \prec Q \sim T$ for all T in $\mathcal{P}_s - \{P, Q\}$ as in part *a*.)
   d. Prove that conditions 1, 2, and 4 imply condition 3.
   e. Argue that conditions 1, 2, and 3 imply condition 4. (See Theorem 3.1.)

7. Show that condition 2 is *not* implied by conditions 1 and 3 of Theorem 8.2 and $C5$ of Theorem 8.3.

8. Show that $A1$, $A2$, and $A3$ of Theorem 8.3 are implied by (8.5) and (8.6).

9. Give details for the assertions in the paragraph following Theorem 8.4.

10. Consider the following two alternatives:

Alternative *A*. One fair coin is flipped. If it lands "heads" you get steak for dinner every night for the next three nights; if it lands "tails" you get chicken for dinner every night for the next three nights.

Alternative *B*. On each of the next three days a fair coin is flipped to determine whether you get steak (if "heads") or chicken (if "tails") for dinner that evening.

Let X be the set of eight triples (x_1, x_2, x_3) where $x_i \in \{\text{chicken, steak}\}$ for $i = 1, 2, 3$ and specify P and Q on X that correspond to alternatives A and B respectively. Can you think of any reasonable argument why $P \sim Q$ ought to be true? Identify your own preference in this case and explain why you prefer the one alternative to the other if you are not indifferent. If you are indifferent, would you remain indifferent if the example were phrased in terms of 100 nights rather than three nights?

11. Consider the following two pairs of gambles in which the

$$\begin{cases} A. & \text{Get \$10 with } pr.\ .3 \text{ or \$50 with } pr.\ .7 \\ B. & \text{Get \$0 with } pr.\ .2 \text{ or \$70 with } pr.\ .8 \end{cases}$$

$$\begin{cases} C. & \text{Get \$20 with } pr.\ .9 \text{ or \$70 with } pr.\ .1 \\ D. & \text{Get \$40 with } pr.\ .6 \text{ or \$60 with } pr.\ .4 \end{cases}$$

amounts of money are to be considered as possible increments to your wealth as of this moment. In considering your preference between A and B the correct interpretation of the expected-utility theory says that you should disregard C and D: that is, suppose you have a choice between A and B and that these are the only two alternatives you can select between and the only two that can change your financial position in the near future. Similarly, disregard A and B when you consider your preference between C and D.

a. Now suppose you are allowed to choose either A or B *and* either C or D before either of your choices is actually played out. You then have four alternatives, say (A, C), (A, D), (B, C), and (B, D). For each of these four alternatives specify the corresponding measure on amounts you might win. Does the theory in this chapter imply that if $A \prec B$ and $C \prec D$, as in the preceding paragraph, then (B, D) will be preferred to the other three alternatives in the new situation? Why not?

b. Suppose you can select either A or B and then, after your selection has been played out, you can choose either C or D and have this second choice played out. Show that you have eight strategies in this case, one of which is: (Select A; if \$10 results then choose C and if \$50 results then choose D). Make out a table that identifies the eight strategies and shows the probability measure on totals you might win with each strategy.

12. Let $x = 0$ represent your present wealth. If P is a probability measure on amounts of money that represent potential incremental additions to your present wealth and if $P \sim \$x$ (where $\$x$ is considered as a sure-thing addition to your present wealth) then $\$x$ is a *certainty equivalent* for P. $P \sim \$x$ means that you would

be indifferent between gambling with P and "receiving" $\$x$ as an outright gift. Estimate your certainty equivalent for P when (*a*) $P(\$0) = .5$, $P(\$10000) = .5$; (*b*) $P(\$0) = .1$, $P(\$1{,}000{,}000) = .9$; (*c*) $P(-\$500) = .5$, $P(\$500) = .5$; (*d*) $P(-\$100) = .2$, $P(-\$10) = .8$; (*e*) $P(\$0) = 1/3$, $P(\$1000) = 1/3$, $P(\$3000) = 1/3$; (*f*) $P(\$90000) = .5$, $P(\$100000) = .5$.

13. (*Continuation.*) Estimate your certainty equivalent for each of the following probability measures.

a. $P(\$0) = .01$, $P(\$5000) = .99$. $\$x \sim P$

b. $Q(\$0) = .99$, $Q(\$5000) = .01$. $\$y \sim Q$

c. $R(\$0) = .50$, $R(\$5000) = .50$. $\$z \sim R$.

Show that the expected-utility theory implies that $R \sim \frac{1}{2}P + \frac{1}{2}Q$. Does this mean that $\$z = \frac{1}{2}(\$x + \$y)$? Does it mean that $\$z$ is indifferent to a 50-50 gamble between $\$x$ and $\$y$?

14. Let $X = \text{Re}$, let u satisfy (8.3) with $x < y$ implying $x \prec y$ and with u on X continuous. Pfanzagl (1959) considers an axiom which when translated into this context reads as follows: *If* $P(x + y) = Q(x)$ *for all* $x \in X$ *and if* $Q \sim z$ *with* $z \in X$ *then* $P \sim y + z$. [Thus, if $P(x + y) = Q(x)$ for all $x \in X$ and if z is the certainty equivalent for Q then $y + z$ is the certainty equivalent for P.]

a. Under the stated conditions Pfanzagl shows that u on X must have one of the following three forms (unique up to a positive linear transformation):
1. $u(x) = k^x$ with $k > 1$, or
2. $u(x) = -k^x$ with $0 < k < 1$, or
3. $u(x) = x$.

Show that each of these expressions satisfies the axiom stated above. Plot (1) with $k = 2$, plot (2) with $k = \frac{1}{2}$, and plot 3.

b. Comment on whether you think this axiom is valid for you. (Consider, for example, your answers to parts *a* and *f* of Exercise 12.)

15. (*Continuation.*) Let $X = \text{Re}$ and let the other conditions in the first sentence of Exercise 14 hold. Show that u on X is *linear* [i.e., case (3) in Exercise 14] if either (*a*) or (*b*) as follows holds with $x \neq y$:

a. For all $x, y \in X$ and all $\alpha \in [0, 1]$, $Q \sim P$ whenever $Q(px + (1 - p)y) = 1$ and $P(x) = p$, $P(y) = 1 - p$.

b. For all $x, y \in X$, $Q \sim P$ whenever $Q((x + y)/2) = 1$ and $P(x) = P(y) = \frac{1}{2}$.

c. Give a critique of these conditions.

16. A man estimates his present wealth at \$50000. Let $x = 0$ correspond to his present wealth and consider possible changes of amounts $\$10000x$ in his present wealth, as shown on Figure 8.4, where $u(x)$ is plotted. For example, $x = 2$ represents an addition of \$20000 to his present wealth. We assume that u has been measured in accord with the expected-utility model. Let A be a 50-50 gamble that pays either \$0 or \$40000.

a. Use Figure 8.4 to estimate the certainty equivalent of A (see Exercise 12). Write out the indifference statement that defines the certainty equivalent in

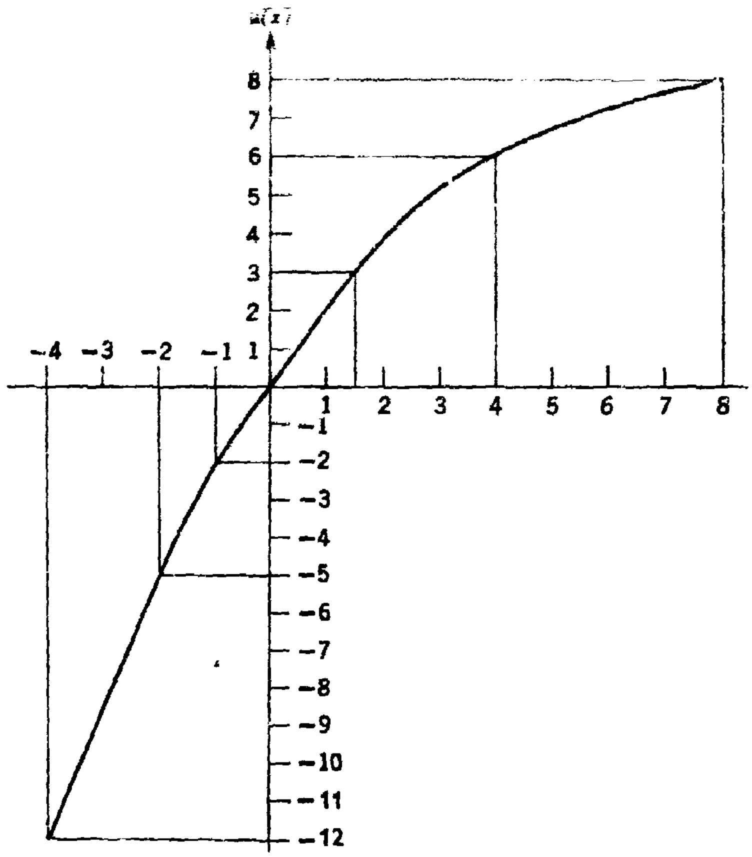

Figure 8.4 Utility function for possible changes in present wealth: \$10000 x is amount of change (see Exercise 8.16).

terms of changes in present wealth, denoting the certainty equivalent by y. [Answer: $y \sim$ (\$40000 with *pr.* $\frac{1}{2}$ or \$0 with *pr.* $\frac{1}{2}$).]

b. If the man is given A as a gift, what is the least amount he would sell it for? Letting y' denote his minimum selling price, write the indifference statement that defines y', and compare to the answer in (*a*).

c. If, instead of being given A, the man considers buying it, what is the most he would pay for it? Letting z be the most he would pay to take possession of A, write the indifference statement that defines z.

d. Suppose the man actually buys A for the amount specified in the answer to (*c*). Will he then be willing to sell it (before it is played out) for the amount specified in the answer to (*b*)? Why not? What would he be willing to sell it for after buying it?

e. Instead of buying A for the amount specified in (*c*) suppose he gets it at a bargain price, say for \$15000. Having bought A for \$15000, what is the minimum amount he would sell it for? Write the defining indifference statement with w the minimum amount.

f. Suppose the man is given A as a gift. He now is given an opportunity to buy a second gamble, also an even-chance gamble for \$0 or \$40000, before A is played out. What is the most he would be willing to pay for the second gamble? Letting r be the most he would pay, write out the indifference statement that defines r. (Do not make the mistake of asserting that $r \sim y'$.)

g. Suppose the man buys A for \$15000 and is then given an opportunity to buy a second gamble just like A before A is played out. What is the most he would pay for this second gamble? Let s be this amount and write out the indifference statement that defines s.

Chapter 9

EXPECTED UTILITY FOR STRICT PARTIAL ORDERS

This chapter examines the important generalization of expected utility for simple probability measures when indifference on $\mathcal{P}_s$ is not assumed to be transitive. We shall consider the representation

$$P < Q \Rightarrow E(u, P) < E(u, Q), \qquad \text{for all } P, Q \in \mathcal{P}_s \tag{9.1}$$

in the context where X is finite. Aumann (1962) and Kannai (1963) discuss the difficulties that arise when X is infinite and Kannai's paper contains several important theorems for this case.

The utility theory in this chapter is largely due to Aumann (1962). Although he assumes that $\preccurlyeq$ is a quasi order (reflexive, transitive), minor revisions make his work applicable to the case where $<$ is a strict partial order (irreflexive, transitive).

Section 9.1 presents an expected-utility theorem and discusses its conditions. The second section develops a support theorem for convex cones in Re^n. The third section proves the utility theorem with the use of the support theorem.

9.1 AN EXPECTED UTILITY THEOREM

In the following theorem $\mathcal{P}_s$ is the set of simple probability measures on X, as in Chapter 8. $\alpha P + (1 - \alpha)Q$ is the direct linear combination of $P, Q \in \mathcal{P}_s$. $E(u, P) = \sum_X u(x)P(x)$.

THEOREM 9.1. *Suppose that X is a finite set and that the following hold throughout $\mathcal{P}_s$ for a binary relation $<$ on $\mathcal{P}_s$:*

1. *$<$ is transitive,*
2. *If $0 < \alpha < 1$ then $P < Q \Leftrightarrow \alpha P + (1 - \alpha)R < \alpha Q + (1 - \alpha)R$,*
3. *If $\alpha P + (1 - \alpha)R < \alpha Q + (1 - \alpha)S$ for all $\alpha \in (0, 1]$ then not $S < R$.*

Then there is a real-valued function u on X that satisfies (9.1).

The three conditions in this theorem compare with the three conditions of Theorem 8.2. The $\Rightarrow$ part of condition 2 in Theorem 9.1 is condition 2 of Theorem 8.2. The $\Leftarrow$ part of condition 2, which is implied by the conditions of Theorem 8.2, can be defended as follows. Suppose in fact that with $\alpha \in (0, 1)$ you prefer $\alpha Q + (1 - \alpha)R$ to $\alpha P + (1 - \alpha)R$. Then it seems reasonable that this preference would depend on your feelings between P and Q. In fact, since the presence of $(1 - \alpha)R$ tends to weaken the difference between the two mixtures, the removal of $(1 - \alpha)R$ should make the distinction between P and Q even clearer than that between $\alpha P + (1 - \alpha)R$ and $\alpha Q + (1 - \alpha)R$ and hence it would seem reasonable that you would prefer Q to P. In the presence of the $\Rightarrow$ part of condition 2 the $\Leftarrow$ part can be written as $[\alpha \in (0, 1), \alpha P + (1 - \alpha)R < \alpha Q + (1 - \alpha)R] \Rightarrow \text{not } P \sim Q$.

The Archimedean axiom, condition 3, is slightly different than Aumann's axiom, which says that if $R < \alpha Q + (1 - \alpha)S$ for all $\alpha \in (0, 1]$ then not $S < R$. However, both axioms are necessary for (9.1). For example, if $\alpha E(u, P) + (1 - \alpha)E(u, R) < \alpha E(u, Q) + (1 - \alpha)E(u, S)$ for all $\alpha \in (0, 1]$, then we cannot have $E(u, S) < E(u, R)$. Therefore, condition 3 is the "weakest" sort of Archimedean condition that can be used to obtain (9.1).

We note also that condition 3 implies that $<$ is irreflexive, and for this reason irreflexivity does not need to be included along with transitivity in condition 1.

Because indifference ($P \sim Q \Leftrightarrow \text{not } P < Q$ and not $Q < P$) is not assumed to be transitive for Theorem 9.1, it is not true in general that u satisfying (9.1) is unique up to a positive linear transformation.

9.2 CONVEX SETS AND CONES

This section develops a theorem from which we shall be able to prove Theorem 9.1. The new theorem states that if a convex cone C in Re^n satisfies specified conditions then there is a $w \in \mathrm{Re}^n$ such that $w \cdot x > 0$ for all $x \in C$. We shall begin with some definitions and two well-known lemmas. In what follows, $X \neq \varnothing$.

A set $X \subseteq \mathrm{Re}^n$ is *convex* if and only if $\alpha x + (1 - \alpha)y \in X$ whenever $x, y \in X$ and $0 \leq \alpha \leq 1$. The *closure* of a convex set X, denoted by $\bar{X}$ as in Section 5.3, is easily shown to be convex also. The topology with respect to which closure is defined is the usual product topology $\mathcal{U}^n$ (Sections 3.4, 5.3). The $\mathbf{0}$ is the origin of Re^n.

LEMMA 9.1. *If* $X \subseteq \mathrm{Re}^n$ *is convex and* $(y \in \mathrm{Re}^n, y \notin \bar{X})$ *then there is a* $w \neq \mathbf{0}$ *in* Re^n *such that*

$$\inf\{w \cdot x : x \in X\} > w \cdot y.$$

Proof. As in the proof of Theorem 4.2 let $z^2 = z \cdot z$. Since $y \notin \bar{X}$, $\inf\{(x - y)^2 : x \in \bar{X}\} > 0$, and it follows from the definitions of closure and convexity that there is a $z \in \bar{X}$ such that $(z - y)^2 = \inf\{(x - y)^2 : x \in \bar{X}\}$. Let $w = z - y$. Hence $w \neq \mathbf{0}$. With $0 < \lambda < 1$ and $x \in \bar{X}$, $(1 - \lambda)z + \lambda x \in \bar{X}$. Hence $((1 - \lambda)z + \lambda x - y)^2 \geq (z - y)^2$. This reduces to $\lambda(x - z)^2 + 2(z - y) \cdot (x - z) \geq 0$. Letting λ approach zero it follows that $w \cdot x \geq w \cdot z$. Since $(z - y) \cdot (z - y) > 0$, $w \cdot z > w \cdot y$. Hence $\inf\{w \cdot x : x \in \bar{X}\} \geq w \cdot z > w \cdot y$. ◆

y is on the *boundary* of X if every open set that contains y contains a point in X and a point not in X.

LEMMA 9.2. *If $X \subseteq \mathrm{Re}^n$ is convex and $y \in \mathrm{Re}^n$ is on the boundary of X then there is a $w \neq \mathbf{0}$ in* Re^n *such that*

$$\inf\{w \cdot x : x \in X\} = w \cdot y. \tag{9.2}$$

Proof. Let $y \in \mathrm{Re}^n$ be on the boundary of convex $X \subseteq \mathrm{Re}^n$. Let Y be an open n-dimensional rectangle that contains y and suppose that $z \in \bar{X}$ for all $z \in Y$. Then, by selecting open rectangles included in Y near the corners of Y, each of these must contain an element in X, and it follows from convexity that there is an open set included in X that contains y. But then y is not on the boundary of X. Hence, for every such Y, there is a point in Y that is not in $\bar{X}$. It then follows that there is a sequence $y_1, y_2, \ldots$ of elements in Re^n that are not in $\bar{X}$ but approach y. Then, by Lemma 9.1, there is a sequence $w_1, w_2, \ldots$ of elements in Re^n that differ from $\mathbf{0}$, have $w_j^2 = 1$ for all j (after multiplication by an appropriate positive number), and satisfy $\inf\{w_j \cdot x : x \in X\} > w_j \cdot y_j$ for $j = 1, 2, \ldots$. Because $w_j^2 = 1$ for all j there must be a $w \in \mathrm{Re}^n$ such that every open n-dimensional rectangle that contains w contains some w_j. It follows that, for each $x \in X$, $w \cdot x \geq w \cdot y$. $\inf\{w \cdot x : x \in X\} > w \cdot y$ is impossible, for if this were so then $w \cdot y > w \cdot y$. ◆

Cones

A set $X \subseteq \mathrm{Re}^n$ is a *cone* if and only if $\alpha x \in X$ whenever $x \in X$ and $\alpha > 0$. A *convex cone* is a cone that is convex. X is a convex cone if and only if $[x, y \in X; \alpha, \beta > 0] \Rightarrow \alpha x + \beta y \in X$. $\mathbf{0}$ is not necessarily an element in a convex cone.

THEOREM 9.2. *Suppose that C is a nonempty convex cone in* Re^n *and that $\bar{C} \cap (-C) = \varnothing$, where $-C = \{x : -x \in C\}$. Then there is a $w \in \mathrm{Re}^n$ such that*

$$w \cdot x > 0 \qquad \textit{for all } x \in C. \tag{9.3}$$

If $0 \in C$ then $0 \in \bar{C}$ and $0 \in -C$ so that $\bar{C} \cap (-C) \neq \varnothing$. Hence the Archimedean condition $\bar{C} \cap (-C) = \varnothing$ requires that $0 \notin C$. This condition is necessary also for (9.3), for if (9.3) holds and $z \in \bar{C} \cap (-C)$, then $w \cdot z < 0$ by (9.3) and hence (since $z \in \bar{C}$) $w \cdot y < 0$ for some $y \in C$.

Proof of Theorem 9.2. The theorem is obviously true when $n = 1$. Using induction we shall assume with $n \geq 2$ that the conclusion follows from the hypotheses for each $m < n$. Thus, let the hypotheses hold for $n \geq 2$. Then, since 0 is on the boundary of C, it follows from Lemma 9.2 that there is a $w \in \text{Re}^n$ with $w \neq 0$ such that

$$w \cdot x \geq 0 \qquad \text{for all } x \in C. \tag{9.4}$$

If (9.3) holds for this w, we are finished. Otherwise $w \cdot z = 0$ for some $z \in C$ and in this case we consider two possibilities.

1. $\{x : w \cdot x > 0\} \subseteq \bar{C}$. Then $\bar{C} = \{x : w \cdot x \geq 0\}$ and with $z \in C$ and $w \cdot z = 0$, $-z \in \bar{C} \cap (-C)$ in violation of the Archimedean condition. Hence this case can't arise under the hypotheses.

2. There is an $x \in \text{Re}^n$ such that $w \cdot x > 0$ and $x \notin \bar{C}$. Let $Y = \{y : x \cdot y = 0\}$. The dimensionality of Y is less than n since $x \neq 0$ and if $x_i \neq 0$ then each $y \in Y$ is uniquely determined by its other $n - 1$ components. Also, each $z \in \text{Re}^n$ is expressible in one and only one way as $\beta x + y$ with $\beta \in \text{Re}$ and $y \in Y$. Namely, $z = (z \cdot x/x^2)x + [z - (z \cdot x/x^2)x]$, and if $z = \beta x + y = \beta' x + y'$ with $\beta \neq \beta'$ then $x = (y - y')/(\beta' - \beta)$, implying $x^2 = x \cdot (y - y')/(\beta' - \beta) = 0$, which is false.

Continuing with Case 2 let

$$C_0 = \{y : \beta x + y \in C \text{ for some } y \in Y \text{ and } \beta \in \text{Re}\}.$$

$C_0 \subseteq Y$ is clearly a nonempty convex cone. To verify that $\bar{C}_0 \cap (-C_0) = \varnothing$ suppose to the contrary that $y \in \bar{C}_0 \cap (-C_0)$. Then there is a $\beta \in \text{Re}$ such that $\beta x - y \in C$ and since $y \in \bar{C}_0$ there is a sequence $y_1, y_2, \ldots$ in C that approaches y $[(y - y_i)^2 \geq (y - y_{i+1})^2$ and $\inf\{(y - y_i)^2 : i = 1, 2, \ldots\} = 0]$ and a sequence of numbers $\beta_1, \beta_2, \ldots$ such that $\beta_i x + y_i \in C$ for all i.

Then $(\beta + \beta_i)x + y_i - y \in C$ for all i so that $(\beta + \beta_i)w \cdot x + w \cdot (y_i - y) \geq 0$ for all i by (9.4): hence the β_i must be bounded below. The β_i must be bounded above also: otherwise there are $x + (y_i/\beta_i) \in C$ that are arbitrarily close to x, and this contradicts $x \notin \bar{C}$. It follows that there is a $\lambda \in \text{Re}$ such that $\inf\{|\lambda - \beta_i| : i = 1, 2, \ldots\} = 0$, and since $\beta_i x + y_i \in C$ for all i it follows that $\lambda x + y \in \bar{C}$. But then $(\lambda x + y) + (\beta x - y) = (\lambda + \beta)x \in \bar{C}$, which is false unless $\lambda + \beta = 0$. But if $\lambda + \beta = 0$ then $\lambda x + y \in \bar{C}$ and $-\lambda x - y \in C$, contradicting $\bar{C} \cap (-C) = \varnothing$.

Therefore $\bar{C}_0 \cap (-C_0) = \varnothing$. It follows from the induction hypothesis for $m < n$ that there is a $v \in Y$ such that $v \cdot y > 0$ for every $y \in C_0$. Since $v \in Y$,

$v \cdot x = 0$ and therefore, for each $z \in C$ written as $z = \beta x + y$ in the C_0 format, $v \cdot z = v \cdot (\beta x + y) = v \cdot y > 0$. ◆

9.3 PROOF OF THEOREM 9.1

Throughout this section the hypotheses of Theorem 9.1 are assumed to hold along with $P < Q$ for some $P, Q \in \mathcal{P}_s$, for otherwise the conclusion is obvious.

Let X have $n + 1$ elements, $n \geq 1$, identified as $x_1, x_2, \ldots, x_{n+1}$. For each $P \in \mathcal{P}_s$ let $p_i = P(x_i)$. Let $\mathcal{P} = \{p = (p_1, \ldots, p_n): p_i \geq 0$ for each i and $\sum p_i \leq 1\}$. Then there is a one-to-one correspondence between $\mathcal{P}_s$ and $\mathcal{P} \subseteq \mathrm{Re}^n$. In terms of $\mathcal{P}$ the conditions are:

1. $(p < q, q < r) \Rightarrow p < r$,
2. If $0 < \alpha < 1$ then $p < q \Leftrightarrow \alpha p + (1 - \alpha)r < \alpha q + (1 - \alpha)r$,
3. If $\alpha p + (1 - \alpha)r < \alpha q + (1 - \alpha)s$ for all $\alpha \in (0, 1]$ then not $s < r$.

Define $D \in \mathrm{Re}^n$ by

$$D = \{t : t = p - q \text{ for some } p, q \in \mathcal{P} \text{ for which } q < p\}. \tag{9.5}$$

Clearly, (9.1) holds if and only if there is a $w \in \mathrm{Re}^n$ such that $w \cdot t > 0$ for every $t \in D$. Some facts about D follow.

a. Suppose $t \in D$ is such that $t = p - q = r - s$ with $q < p$. Then $\frac{1}{2}r + \frac{1}{2}q = \frac{1}{2}p + \frac{1}{2}s$, $\frac{1}{2}r + \frac{1}{2}q < \frac{1}{2}r + \frac{1}{2}p$ by condition 2, and therefore $\frac{1}{2}p + \frac{1}{2}s < \frac{1}{2}r + \frac{1}{2}p$, so that $s < r$ by condition 2 ($\Leftarrow$). Hence *if $q < p$ then $s < r$ whenever $r - s = p - q$.*

b. Suppose $t = p - q$ and $t^* = r - s$ are in D. Then $q < p$ and $s < r$. Hence, by conditions 1 and 2, $\alpha q + (1 - \alpha)s < \alpha p + (1 - \alpha)r$ for any $\alpha \in (0, 1)$, and hence $\alpha t + (1 - \alpha)t^* \in D$. Thus *$D$ is convex.*

c. *If $t = p - q$ for some $p, q \in \mathcal{P}$ then $t \in D \Leftrightarrow \alpha t \in D$ for all $\alpha \in (0, 1)$.* This follows from condition 2.

d. *If $t = q - p$ and $t^* = s - r$ for $p, q, r, s \in \mathcal{P}$ and if $\alpha t + (1 - \alpha)t^* \in D$ for all $\alpha \in (0, 1]$ then $-t^* \notin D$.* To prove this observe that $\alpha t + (1 - \alpha)t^* \in D$ implies that $\alpha p + (1 - \alpha)r < \alpha q + (1 - \alpha)s$ by (*a*). Then, by condition 3, not $s < r$. Hence, again using (*a*) with $-t^* = r - s$, $-t^* \notin D$.

Based on D we define a cone C as follows:

$$C = \{x : x = \alpha t \text{ for some } \alpha > 0 \text{ and } t \in D\}.$$

Since $D \neq \varnothing$ by assumption, $C \neq \varnothing$. The convexity of C follows easily from properties (*b*) and (*c*) for D. For the Archimedean condition we wish to have:

$$\alpha t + (1 - \alpha)t^* \in C \quad \text{for all } \alpha \in (0, 1] \Rightarrow -t^* \notin C. \tag{9.6}$$

This is obviously true if $t^* = 0$. Henceforth take $t^* \neq 0$. If $t \in C$ it is easily seen (Exercise 11) that there is a $\beta > 0$ such that $\beta t \in D$ and $(\beta t)^2 \geq 1/n^2$. Given $\alpha t + (1 - \alpha)t^* \in C$ for all $\alpha \in (0, 1]$, it follows that for each $\alpha \in (0, 1]$ there is a $\beta(\alpha) > 0$ such that $\alpha(\beta(\alpha)t) + (1 - \alpha)(\beta(\alpha)t^*) \in D$ and $\beta(\alpha)^2(\alpha t + (1 - \alpha)t^*)^2 \geq 1/n^2$. Since $\{(\alpha t + (1 - \alpha)t^*)^2 : \alpha \in (0, 1]\}$ is bounded above, it follows that $\beta(a) > \delta$ for some $\delta > 0$ and all $\alpha \in (0, 1]$. Therefore there is a $\beta > 0$ such that $\alpha(\beta t) + (1 - \alpha)(\beta t^*) \in D$ for all $\alpha \in (0, 1]$. With β such that $(\beta t^*)^2 \leq 1/n^2$ it follows from (*d*) that $-\beta t^* \notin D$ and hence that $-t^* \notin C$. This verifies (9.6).

Suppose $C \subseteq \mathrm{Re}^n$ is actually n dimensional so that some $t \in C$ is not on the boundary of C. Then there is an open n-dimensional cube in Re^n that contains such a t and is included in C. It follows with little difficulty that if $z \in \bar{C}$ then $\alpha t + (1 - \alpha)z \in C$ for all $\alpha \in (0, 1]$, and hence $-z \notin C$ by (9.6). Hence $\bar{C} \cap (-C) = \varnothing$ and it follows from Theorem 9.2 that there is a $w \in \mathrm{Re}^n$ such that $w \cdot t > 0$ for all $t \in C$ and hence $w \cdot t > 0$ for all $t \in D$. If every point in C is on the boundary of C (with respect to Re^n) then the dimensionality of C is less than n and a similar analysis applies with respect to the actual dimensionality of C. ◆

9.4 SUMMARY

$C \subseteq \mathrm{Re}^n$ is a convex cone if $[x, y \in C; \alpha, \beta > 0] \Rightarrow \alpha x + \beta y \in C$. If C is a nonempty convex cone in Re^n, and $-C$ and the closure of C have no point in common, then there is a $w \in \mathrm{Re}^n$ such that $w_1x_1 + \cdots + w_nx_n > 0$ for every x in C. This result can be used to prove that if X is finite, and if $<$ on $\mathcal{P}_s$ is a strict partial order that satisfies an appropriate independence condition and a necessary Archimedean condition, then there is a real-valued function u on X that satisfies $P < Q \Rightarrow E(u, P) < E(u, Q)$ for all P and Q in $\mathcal{P}_s$.

INDEX TO EXERCISES

1. Independence conditions. 2–3. Conditions that imply transitive indifference. 4. Aumann's theorem. 5. $P \approx Q \Leftrightarrow P = Q$. 6–7. Convex sets and closure. 8. Limit point. 9–10. More boundaries. 11–12. Distance from the origin. 13. Linear additivity.

Exercises

1. With $\prec$ on $\mathcal{P}_s$ a strict partial order let $P \sim Q \Leftrightarrow$ (not $P \prec Q$, not $Q \prec P$) and $P \approx Q \Leftrightarrow (P \sim R \Leftrightarrow Q \sim R$, for all $R \in \mathcal{P}_s$), as usual. Let $B1$, $B2$, and $B3$ be the

following independence conditions:

$B1$. $[P \prec Q, 0 < \alpha < 1] \Rightarrow \alpha P + (1 - \alpha)R \prec \alpha Q + (1 - \alpha)R.$
$B2$. $[P \sim Q, 0 < \alpha < 1] \Rightarrow \alpha P + (1 - \alpha)R \sim \alpha Q + (1 - \alpha)R.$
$B3$. $[P \approx Q, 0 < \alpha < 1] \Rightarrow \alpha P + (1 - \alpha)R \approx \alpha Q + (1 - \alpha)R.$

Express your opinions on the reasonableness of $B2$ and $B3$, show that $(B1, B2)$ imply the converse $(\Leftarrow)$ of each of $B1$, $B2$, and $B3$ (with $0 < \alpha < 1$), and construct a specific example to show that $B1$, $B2$, and $B3$ do not imply that $\sim$ is transitive. Assume throughout that $\prec$ on $\mathcal{P}_s$ is a strict partial order.

2. (*Continuation.*) Let $C1$ and $C2$ be respectively the semiorder conditions $(P \prec Q, Q \prec R) \Rightarrow (P \prec S \text{ or } S \prec R)$ and $(P \prec Q, R \prec S) \Rightarrow (P \prec S \text{ or } R \prec Q)$. Assume that $\prec$ is irreflexive.

a. Construct situations that question the reasonableness of $C1$ and $C2$.
b. Show that $(B1, C1) \Rightarrow \sim$ is transitive.
c. Show that $(B1, B2, C2) \Rightarrow \sim$ is transitive.

3. (*Continuation.*) Let $B4$ be the condition: $(P \sim R, Q \sim R, 0 < \alpha < 1) \Rightarrow \alpha P + (1 - \alpha)Q \sim R$. Show that $B4$ is implied by strict partial order and $B1$ provided that $P \prec Q$ or $Q \prec P$. Then prove that (strict partial order, $B1$, $B2$, $B4$) $\Rightarrow \sim$ is transitive. Can you construct a situation that questions the reasonableness of $B4$? If so, what is it?

4. Aumann (1962) proves that if $\preccurlyeq^*$ on $\mathcal{P}_s$ for finite X is a quasi order (reflexive, transitive), if $P \preccurlyeq^* Q \Leftrightarrow \alpha P + (1 - \alpha)R \preccurlyeq^* \alpha Q + (1 - \alpha)R$ whenever $0 < \alpha < 1$, and if $R \prec^* \alpha P + (1 - \alpha)Q$ for all $\alpha \in (0, 1] \Rightarrow$ not $Q \prec^* R$, then there is a real-valued function u on X such that, for all $P, Q \in \mathcal{P}_s$, $P \prec^* Q \Rightarrow E(u, P) < E(u, Q)$ and $P \sim^* Q \Rightarrow E(u, P) = E(u, Q)$. Here $P \prec^* Q \Leftrightarrow (P \preccurlyeq^* Q,$ not $Q \preccurlyeq^* P)$ and $P \sim^* Q \Leftrightarrow (P \preccurlyeq^* Q, Q \preccurlyeq^* P)$. Now assume that $\prec$ on $\mathcal{P}_s$ is a strict partial order that satisfies $B1$, $B2$, and $B3$ of Exercise 1 along with $R \prec \alpha P + (1 - \alpha)Q$ for all $\alpha \in (0, 1] \Rightarrow$ not $Q \prec R$. Defining $\preccurlyeq^*$ from $\prec$ by $P \preccurlyeq^* Q \Leftrightarrow (P \prec Q \text{ or } P \approx Q)$, show that $\preccurlyeq^*$ satisfies Aumann's conditions and hence that there is a real-valued function u on X (finite) such that (9.1) holds along with

$$P \approx Q \Rightarrow E(u, P) = E(u, Q), \qquad \text{for all} \quad P, Q \in \mathcal{P}_s.$$

5. Suppose $X = \{\$1, \$2, \ldots, \$100\}$, with $\$1 \prec \$2 \prec \cdots \prec \$100$. Argue that with $\approx$ defined as in Exercise 1, it would not be unusual to find that $P \approx Q \Leftrightarrow P = Q$ when $\prec$ on $\mathcal{P}_s$ is a strict partial order. Can you think of a case (with elements in X not monetary) where $P \approx Q$ would seem reasonable for some P, Q with $P \neq Q$?

6. Prove that if $X \subseteq \mathrm{Re}^n$ is convex then so is $\bar{X}$.

7. Show that if $X \subseteq \mathrm{Re}^n$ is convex and $(y \in \mathrm{Re}^n, y \notin \bar{X})$ then there is a $z \in \bar{X}$ such that $(z - y)^2 = \inf\{(x - y)^2 : x \in \bar{X}\}$.

8. Let $w_j \in \mathrm{Re}^n$ be such that $w_j^2 = 1$ for $j = 1, 2, \ldots$. Prove that there is a $w \in \mathrm{Re}^n$ such that every open n-dimensional rectangle that contains w contains some w_j.

9. Describe the boundaries of the following convex sets in Re^2: (*a*) $\{x : x_1^2 + x_2^2 \leq 1\}$, (*b*) $\{x : x_1^2 + x_2^2 < 1\}$, (*c*) $\{x : 0 < x_1 < 1, 0 \leq x_2 \leq 1\}$, and (*d*) $\{x : x = (\alpha, \alpha)$ for $\alpha \in (0, 1]\}$.

10. With X a convex set in Re^n suppose that $t \in X$ is not on the boundary of X. Verify that if $z \in \bar{X}$ then $\alpha t + (1 - \alpha)z \in X$ for all $\alpha \in (0, 1]$.

11. With D as defined by (9.5) suppose $t \in D$. Define $p, q \in \mathcal{T}$ as follows: $(p_i, q_i) = (t_i, 0)$ or $(0, t_i)$ or $(0, 0)$ according to whether $t_i > 0$ or $t_i < 0$ or $t_i = 0$. Then $p - q = t$. Now multiply every t_i by $\alpha > 0$ with α as large as possible so that $\sum_{\{i: t_i > 0\}} \alpha t_i \leq 1$ and $\sum_{\{i: t_i < 0\}} \alpha t_i \geq -1$. Then αp, $\alpha q \in \mathcal{T}$ and $\alpha t = \alpha p - \alpha q$ is in D. Verify that $\sum_{i=1}^n (\alpha t_i)^2 \geq 1/n^2$.

12. (*Continuation.*) Verify that $\sum_{i=1}^n (\alpha t_i)^2 \geq 1/n$.

13. Argue from the theory in this chapter that if X is the non-negative orthant of Re^n and if, for all $x, y, z, w \in X$,

a. $\prec$ is transitive,

b. If $\alpha \in (0, 1)$ then $x \prec y \Leftrightarrow \alpha x + (1 - \alpha)z \prec \alpha y + (1 - \alpha)z$,

c. $\alpha x + (1 - \alpha)y \prec \alpha z + (1 - \alpha)w$ for all $\alpha \in (0, 1] \Rightarrow$ not $w \prec y$, then there are real numbers $\lambda_1, \ldots, \lambda_n$ such that $x \prec y \Rightarrow \sum_{i=1}^n \lambda_i x_i < \sum_{i=1}^n \lambda_i y_i$, for all $x, y \in X$. What must be true of the λ_i if (1) ($x_i \leq y_i$ for all i, $x \neq y$) $\Rightarrow x \prec y$, (2) ($x_i < y_i$ for all i) $\Rightarrow x \prec y$?

Chapter 10

EXPECTED UTILITY FOR PROBABILITY MEASURES

This chapter extends the weak order expected-utility theory of Chapter 8 to more general sets of probability measures. Since the sets of measures considered are mixture sets, Theorem 8.4 will be used as a base for establishing the representation $P < Q \Leftrightarrow E(u, P) < E(u, Q)$. Conditions that go beyond those of Theorem 8.4 are required for the extensions. The primary new condition says that if a measure P is preferred to every consequence in a subset Y of consequences for which $Q(Y) = 1$, then Q shall not be preferred to P.

After two preliminary examples, Sections 10.2 and 10.3 develop necessary background material on probability measures and expectations. The actual utility theory development begins in Section 10.4.

10.1 TWO EXAMPLES

In our first example, a decision maker must decide between two construction procedures, A and B, for building a bridge over a river. Procedure A will cost \$150 million and B will cost \$100 million. For A engineers have estimated the probability $P(t)$ of completing the bridge by t years from now at 0 for $t \leq 2$ and $(t - 2)/3$ for $2 \leq t \leq 5$. For B, the probability $Q(t)$ of completion by t years from now is estimated at 0 for $t \leq 3$ and $(t - 3)/4$ for $3 \leq t \leq 7$.

The decision maker's utilities for the applicable consequences are estimated according to the expected-utility model as $u(\$150, t) = -(t - 2)^2 - 5$ for procedure A, and as $u(\$100, t) = -(t - 2)^2$ for procedure B. The expected utility of A is therefore

$$\int_2^5 [-(t - 2)^2 - 5](1/3)\, dt = -8$$

and for B the expected utility is

$$\int_3^7 [-(t-2)^2](1/4)\,dt = -10.33.$$

Thus procedure A, more costly but faster than B, has the greater expected utility.

The St. Petersburg Game

The often-discussed "St. Petersburg game" from Bernoulli (1738) gives an example of a discrete probability measure. Consider a sequence of coin tosses and let α_n be the probability that a "head" occurs for the first time at the nth toss. Suppose you believe that $\alpha_n = 2^{-n}$ for $n = 1, 2, \ldots$ and are given a choice between "Don't play" and "Pay the house \$100 and get back \$$2^n$ if the first head occurs at the nth toss."

Let X be amounts of money representing changes in your present wealth. Then, with u defined on X,

Expected utility of "Don't play" $= u(\$0)$

Expected utility of "Pay and play" $= \sum_{n=1}^{\infty} u(\$2^n - \$100)2^{-n}$.

According to the theory given later, u on X is bounded. Suppose, for example, that $u(x) = x/(|x| + 10000)$, so that $-1 < u(x) < 1$ for all x. Then $u(\$0) = 0$ and $\sum u(\$2^n - \$100)2^{-n} < 0$ so that "Don't play" has the greater expected utility.

10.2 PROBABILITY MEASURES

Generally speaking, probability measures are defined on Boolean algebras of sets. In the following definition $A^c = \{x : x \in X, x \notin A\}$, the *complement* of A with respect to X, and

$$\bigcup_{i=1}^{\infty} A_i = \{x : x \in A_i \quad \text{for some } i \in \{1, 2, \ldots\}\}$$

Definition 10.1. A *Boolean algebra* $\mathcal{A}$ for X is a set of subsets of X such that

1. $X \in \mathcal{A}$,
2. $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$,
3. $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$.

A *σ-algebra* $\mathcal{A}$ for X is a Boolean algebra that satisfies

4. $A_i \in \mathcal{A}$ for $i = 1, 2, \ldots \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

$\{\varnothing, X\}$ is the smallest Boolean and σ-algebra for nonempty X. The largest

Boolean and σ-algebra is the set of all subsets of X. For reasons that will become clearer later we shall usually assume that $\{x\} \in \mathcal{A}$ for each $x \in X$.

If X is finite then every Boolean algebra is a σ-algebra. The difference between these two arises when X is infinite and it has some affect on properties of probability measures. Some authors, such as Loève (1960), deal exclusively with σ-algebras (or "σ-fields").

If $\mathcal{C}$ is an arbitrary set of subsets of X, the *Boolean algebra generated by* $\mathcal{C}$ (minimal Boolean algebra over $\mathcal{C}$) is the intersection of all Boolean algebras that include $\mathcal{C}$. The *σ-algebra generated by* $\mathcal{C}$ is the intersection of all σ-algebras that include $\mathcal{C}$. It is easily verified that the intersection of a set of Boolean (σ) algebras for X is a Boolean (σ) algebra for X.

With $X = \{1, 2, \ldots\}$ and $\mathcal{C} = \{\{1\}, \{2\}, \ldots\}$, the set of all unit subsets of X, the Boolean algebra $\mathcal{M}$ generated by $\mathcal{C}$ is the set of all subsets of X that are either finite or contain all but a finite number of elements in X. But $\mathcal{M}$ is not a σ-algebra since it doesn't contain the set of all even, positive integers. The σ-algebra generated by $\mathcal{C}$ is the set of all subsets of X.

Let $X = \text{Re}$, with $\mathcal{C}$ the set of all intervals in Re. The σ-algebra generated by $\mathcal{C}$ is called the *Borel algebra* for Re, and its elements are *Borel sets*. There are subsets of Re that are not Borel sets: see, for example, Halmos (1950, pp. 66–72).

Throughout the rest of this chapter, $\mathcal{A}$ denotes an algebra (Boolean or σ) for X.

Probability Measures and Countable Convex Combinations

Definition 10.2. A *probability measure* on $\mathcal{A}$ is a real-valued function P on $\mathcal{A}$ such that

1. $P(A) \geq 0$ for every $A \in \mathcal{A}$,
2. $P(X) = 1$,
3. $[A, B \in \mathcal{A}, A \cap B = \varnothing] \Rightarrow P(A \cup B) = P(A) + P(B)$.

For further definitions we shall use the standard notation

$$\sum_{i=1}^{\infty} p_i = \sup\left\{\sum_{i=1}^{n} p_i : n = 1, 2, \ldots\right\} \tag{10.1}$$

when $p_i \geq 0$ for all i and $\sum_{i=1}^{\infty} p_i \leq M$ for some M and all $n = 1, 2, \ldots$. Since $\sum_{i=1}^{n} 2^{-i} = 1 - 2^{-n}$, $\sum_{i=1}^{\infty} 2^{-i} = 1$.

Definition 10.3. If P_i is a probability measure on $\mathcal{A}$ and $\alpha_i \geq 0$ for $i = 1, 2, \ldots$, and if $\sum_{i=1}^{\infty} \alpha_i = 1$, then $\sum_{i=1}^{\infty} \alpha_i P_i$ is the function on $\mathcal{A}$ that assigns the number $\sum_{i=1}^{\infty} \alpha_i P_i(A)$ to each $A \in \mathcal{A}$.

The proof of the following lemma is left to the reader.

LEMMA 10.1. *$\sum_{i=1}^{\infty} \alpha_i P_i$ as defined in Definition 10.3 is a probability measure on $\mathcal{A}$.*

The next definition will be used in our utility theory.

Definition 10.4. A set $\mathcal{P}$ of probability measures on $\mathcal{A}$ is *closed under countable convex combinations* if and only if $\sum_{i=1}^{\infty} \alpha_i P_i \in \mathcal{P}$ whenever $P_i \in \mathcal{P}$ and $\alpha_i \geq 0$ for $i = 1, 2, \ldots$, and $\sum_{i=1}^{\infty} \alpha_i = 1$.

If $\mathcal{P}$ is closed under countable convex combinations then $\mathcal{P}$ is a mixture set (Definition 8.3). Hence if $<$ on $\mathcal{P}$ satisfies $A1$, $A2$, and $A3$ of Theorem 8.3 then (8.5) and (8.6) hold and u on $\mathcal{P}$ is unique up to a positive linear transformation.

Countably-Additive Probability Measures

Definition 10.5. A probability measure P on $\mathcal{A}$ is *countably additive* if and only if

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) \tag{10.2}$$

whenever $A_i \in \mathcal{A}$ for $i = 1, 2, \ldots$, $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ and $A_i \cap A_j = \varnothing$ when $i \neq j$.

This applies whether $\mathcal{A}$ is a σ-algebra or a Boolean algebra that is not also a σ-algebra. (10.2) is an extension of Definition 10.2 (3).

Let $\mathcal{M}$ be the Boolean algebra generated by $\mathcal{C} = \{\{1\}, \{2\}, \ldots\}$, and let P on $\mathcal{M}$ be defined on the basis of $P(n) = 2^{-n}$ for each $n \in X = \{1, 2, \ldots\}$. P is countably additive but $\mathcal{M}$ is not a σ-algebra.

Let $\mathcal{A}$ be the set of all subsets of $\{1, 2, \ldots\}$ and let P on $\mathcal{A}$ be any probability measure that has $P(n) = 0$ for each $n \in \{1, 2, \ldots\}$. Then $\mathcal{A}$ is a σ-algebra and P is not countably additive. Dubins and Savage (1965) call any measure that assigns probability 1 to a denumerable subset of X and probability 0 to every unit subset *diffuse*. The uniform measure on the positive integers, with $P(n) = 0$ for $n = 1, 2, \ldots$ and $P(\{n, 2n, 3n, \ldots\}) = 1/n$ for $n = 1, 2, \ldots$, is diffuse.

Let $\mathcal{A}$ be the set of all Borel sets in $[0, 1]$, and let P be the uniform measure on $\mathcal{A}$ defined on the basis of $P([a, b]) = b - a$ when $0 \leq a \leq b \leq 1$. This P is a countably-additive measure on a σ-algebra.

An important property of countably-additive measures is noted in the next lemma.

LEMMA 10.2. *If P on $\mathcal{A}$ is countably additive, if $\mathcal{B}$ is a countable subset of $\mathcal{A}$ whose elements are weakly ordered by $\subset$, and if $\bigcup_{\mathcal{B}} A \in \mathcal{A}$ then*

$$P(\textstyle\bigcup_{\mathcal{B}} A) = \sup\{P(A): A \in \mathcal{B}\}. \tag{10.3}$$

Proof. The conclusion is obvious if $\mathcal{B}$ is finite. Assume then that $\mathcal{B}$ is denumerable, enumerated as $A_1, A_2, A_3, \ldots$. Let $C_n = \bigcup_{i=1}^{n} A_i$. Then $C_1 \subseteq C_2 \subseteq C_3 \subseteq \cdots$, $\bigcup_{\mathcal{B}} A = \bigcup_{i=1}^{\infty} C_i$ and $\sup\{P(A): A \in \mathcal{B}\} = \sup\{P(C_n): n = 1, 2, \ldots\}$. This last equality follows from the facts that for

any $A \in \mathcal{B}$ there is an n such that $P(C_n) \geq P(A)$ and that for any C_n there is an $A \in \mathcal{B}$ such that $P(A) \geq P(C_n)$.

Let $D_1 = C_1$ and $D_i = C_i - C_{i-1}$ (set theoretic subtraction) for $i = 2, 3, \ldots$, so that $\bigcup D_i = \bigcup C_i$, $D_i \cap D_j = \varnothing$ whenever $i \neq j$, and $C_n = \bigcup_{i=1}^{n} D_i$. Then

$$\begin{aligned}
P\left(\bigcup_{\mathcal{B}} A\right) &= P\left(\bigcup_{i=1}^{\infty} D_i\right) && \text{since } \bigcup_{\mathcal{B}} A = \bigcup_{i=1}^{\infty} D_i \\
&= \sum_{i=1}^{\infty} P(D_i) && \text{by countable additivity} \\
&= \sup\left\{\sum_{i=1}^{n} P(D_i): n = 1, 2, \ldots\right\} && \text{by definition} \\
&= \sup\{P(C_n): n = 1, 2, \ldots\} && \text{by finite additivity} \\
&= \sup\{P(A): A \in \mathcal{B}\}. \quad \blacklozenge
\end{aligned}$$

Discrete Probability Measures

Definition 10.6. A probability measure P on $\mathcal{A}$ is *discrete* if and only if $\{x\} \in \mathcal{A}$ for each $x \in X$, $\mathcal{A}$ is a σ-algebra, P is countably additive and $P(A) = 1$ for some countable $A \in \mathcal{A}$.

All simple measures are discrete. Nonsimple discrete measures on the set of all subsets of $X = \{0, 1, 2, \ldots\}$ include the geometric distributions $[P(n) = p(1-p)^n, 0 < p < 1]$ and Poisson distributions $[P(n) = e^{-\gamma}\lambda^n/n!, \lambda > 0]$. The following lemma compares with Theorem 8.1.

LEMMA 10.3. *If P on $\mathcal{A}$ is discrete then $P(x) = 0$ for all but a countable number of $x \in X$ and*

$$P(A) = \sum_{x \in A} P(x) \qquad \text{for all } A \in \mathcal{A}. \tag{10.4}$$

Proof. Let A be countable with $P(A) = 1$ Then $P(x) = 0$ for every $x \in A^c$ for otherwise $P(A \cup \{x\}) > 1$ for some $x \in A^c$. (10.4) follows from (10.2) when A is countable. (10.4) holds in general if $P(C) = 0$ when $P(x) = 0$ for all $x \in C$ and $C \in \mathcal{A}$. Let $D = \{x: x \in X, P(x) > 0\}$. If $P(D) < 1$ it follows from (10.4) for countable sets that $P(A) < 1$ for every countable $A \in \mathcal{A}$, a contradiction. Hence $P(D) = 1$. Then $P(C) = 0$ when $C \cap D = \varnothing$. $\blacklozenge$

Lemma 10.3 shows that a discrete measure is completely described by the point probabilities $P(x)$.

Conditional Probability Measures

Definition 10.7. If P on $\mathcal{A}$ is a probability measure and if $A \in \mathcal{A}$ and $P(A) > 0$ then the *conditional measure* of P given A, written P_A, is the

function defined by

$$P_A(B) = P(B \cap A)/P(A) \qquad \text{for all } B \in \mathcal{A}. \tag{10.5}$$

When P_A is well-defined, it is a probability measure on $\mathcal{A}$: if P is countably additive then so is P_A. $P_A(A) = 1$, $P_A(B) = 1$ if $A \subseteq B \in \mathcal{A}$, $P_A(B) = P(B)/P(A)$ if $B \subseteq A$ and $B \in \mathcal{A}$. If $A, B \in \mathcal{A}$ and $P(A) > 0$ and $P(B) > 0$ then

$$P(A)P_A(B) = P(B)P_B(A) = P(A \cap B).$$

$P_A(B)$ can be interpreted as the probability that the consequence that occurs will be in B, given that the consequence that occurs will be in A. If $B \cap A = \varnothing$ then $P_A(B) = 0$.

If $P(A) > 0$ and $P(A^c) > 0$ then (note the convex combination)

$$P = P(A)P_A + P(A^c)P_{A^c} \tag{10.6}$$

since, for any $B \in \mathcal{A}$, $P(B) = P(B \cap (A \cup A^c)) = P((B \cap A) \cup (B \cap A^c)) = P(B \cap A) + P(B \cap A^c) = P(A)P(B \cap A)/P(A) + P(A^c)P(B \cap A^c)/P(A^c) = P(A)P_A(B) + P(A^c)P_{A^c}(B)$. More generally, if $P(A) = 1$ (with $A \in \mathcal{A}$), if $\{A_1, \ldots, A_n\}$ is an $\mathcal{A}$-partition of A, and if $I = \{i : P(A_i) > 0\}$, then

$$P = \sum_I P(A_i)P_{A_i} \tag{10.7}$$

since, for any $B \in \mathcal{A}$,

$$\begin{aligned}
\sum_I P(A_i)P_{A_i}(B) &= \sum_I P(B \cap A_i) && \text{by (10.5)} \\
&= \sum_{i=1}^{n} P(B \cap A_i) \\
&= P\left(\bigcup_{i=1}^{n} B \cap A_i\right) && \text{by finite additivity} \\
&= P(B \cap A) = P(B \cap A) + P(B \cap A^c) = P(B).
\end{aligned}$$

(10.7) holds also when P is countably additive, $A \in \mathcal{A}$ and $P(A) = 1$, and $\{A_1, A_2, \ldots\}$ is a denumerable $\mathcal{A}$-partition of A and $I = \{i : P(A_i) > 0\}$.

The following definition will be used in our utility development.

Definition 10.8. A set $\mathcal{P}$ of probability measures on $\mathcal{A}$ is *closed under the formation of conditional probabilities* if and only if $[P \in \mathcal{P}, A \in \mathcal{A}, P(A) > 0] \Rightarrow P_A \in \mathcal{P}$.

10.3 EXPECTATIONS

This section defines precisely the expected value $E(f, P)$ of a bounded, real-valued function f on X with respect to a probability measure P on $\mathcal{A}$. In general we shall assume that f is $\mathcal{A}$-measurable.

Definition 10.9. f is *$\mathcal{A}$-measurable* if and only if f is a real-valued function on X and $\{x:f(x) \in I\} \in \mathcal{A}$ for every interval $I \in \mathrm{Re}$.

$\mathcal{A}$-measurable functions are sometimes called random variables, but it is more common to use this term for functions f on X for which $\{x:f(x) \in B\} \in \mathcal{A}$ for every Borel set $B \in \mathrm{Re}$.

To define expectation we begin with simple $\mathcal{A}$-measurable functions.

Definition 10.10. An $\mathcal{A}$-measurable function f is *simple* if and only if $\{f(x):x \in X\}$ is finite. If f is simple and takes on n distinct values $c_1, \ldots, c_n$ with $f(x) = c_i$ for all $x \in A_i$ then each $A_i \in \mathcal{A}$ by Definition 10.9 and $\{A_1, \ldots, A_n\}$ is a partition of X: with P a probability measure on $\mathcal{A}$, we then define

$$E(f, P) = \sum_{i=1}^{n} c_i P(A_i). \tag{10.8}$$

Simple $\mathcal{A}$-measurable functions are bounded. In general, an $\mathcal{A}$-measurable function f is *bounded* if and only if there are numbers a and b for which $a \leq f(x) \leq b$ for all $x \in X$. In defining $E(f, P)$ for any bounded, $\mathcal{A}$-measurable f we shall use

Definition 10.11. A sequence $f_1, f_2, \ldots$ of simple $\mathcal{A}$-measurable functions *converges uniformly from below* to an $\mathcal{A}$-measurable function f if and only if, for all $x \in X$,

1. $f_1(x) \leq f_2(x) \leq \cdots$
2. $f(x) = \sup\{f_n(x):n = 1, 2, \ldots\}$
3. For any $\epsilon > 0$ there is a positive integer n (which may depend on ϵ) such that $f(x) \leq f_n(x) + \epsilon$.

For any bounded, $\mathcal{A}$-measurable f there is a sequence of simple $\mathcal{A}$-measurable functions that converges uniformly from below to f. With $X = \{x:x \in X \text{ and } a \leq f(x) \leq b\}$ and f $\mathcal{A}$-measurable let

$$\begin{aligned} A_{1,n} &= \{x:a \leq f(x) \leq a + (b - a)/n\} \\ A_{i,n} &= \{x:a + (i - 1)(b - a)/n < f(x) \leq a + i(b - a)/n\} \qquad i = 2, \ldots n, \end{aligned} \tag{10.9}$$

and define f_n by

$$f_n(x) = a + (i - 1)(b - a)/n \qquad \text{for all } x \in A_{i,n}, \qquad i = 1, \ldots, n. \tag{10.10}$$

Each $A_{i,n} \in \mathcal{A}$ by Definition 10.9 and therefore each f_n is a simple $\mathcal{A}$-measurable function. Conditions 1 and 2 of Definition 10.11 are easily verified and condition 3 holds with $n \geq (b - a)/\epsilon$.

Definition 10.12. If f is bounded and $\mathcal{A}$-measurable and if P is a probability measure on $\mathcal{A}$ then

$$E(f, P) = \sup\{E(f_n, P): n = 1, 2, \ldots\} \tag{10.11}$$

where $f_1, f_2, \ldots$ is any sequence of simple $\mathcal{A}$-measurable functions that converges uniformly from below to f.

The following lemma notes that $E(f, P)$ is well defined.

LEMMA 10.4. *If $f_1, f_2, \ldots$ and $g_1, g_2, \ldots$ are sequences of simple $\mathcal{A}$-measurable functions that converge uniformly from below to a bounded, $\mathcal{A}$-measurable function f then* $\sup\{E(f_n, P): n = 1, 2, \ldots\}$ *is finite and*

$$\sup\{E(f_n, P): n = 1, 2, \ldots\} = \sup\{E(g_n, P): n = 1, 2, \ldots\}. \tag{10.12}$$

Proof. Boundedness assures a finite sup. To verify (10.12) assume to the contrary that $\sup\{E(f_n, P): n = 1, 2, \ldots\} < \sup\{E(g_n, P): n = 1, 2, \ldots\}$. Then there is an $\epsilon > 0$ and a positive integer m such that

$$E(f_n, P) + \epsilon < E(g_m, P) \qquad \text{for } n = 1, 2, \ldots. \tag{10.13}$$

By condition 3 of Definition 10.11 there is a k such that $f(x) \leq f_k(x) + \epsilon$ for all $x \in X$, so that $E(h, P) \leq E(f_k + \epsilon, P)$ for every simple $\mathcal{A}$-measurable h for which $h(x) \leq f(x)$ for all $x \in X$. In particular, $E(g_m, P) \leq E(f_k + \epsilon, P) = E(f_k, P) + \epsilon$, contradicting (10.13). ◆

$E(f, P_A)$ for a well-defined conditional probability measure P_A is defined as above since P_A is a probability measure on $\mathcal{A}$.

Finite versus Countable Additivity

The uniformity condition 3 of Definition 10.11 is superfluous for defining $E(f, P)$ when P is countably additive and $\mathcal{A}$ is a σ-algebra (Exercise 15). But uniform convergence is required when countable additivity is not assumed to hold. The following illustrates what amounts to the failure of (10.3) for a diffuse measure.

Let $X = \{0, 1, 2, \ldots\}$, let $\mathcal{A}$ be the set of all subsets of X (a σ-algebra), let P be any probability measure on $\mathcal{A}$ that has $P(x) = 0$ for all $x \in X$, and let $f(x) = x/(1 + x)$ for all $x \in X$.

Since $0 \leq f(x) < 1$ on X we can let $a = 0$ and $b = 1$ in (10.9) and (10.10) to obtain $E(f_n, P) = \sum_{i=1}^{n} [(i - 1)/n]P(A_{i,n}) = (n - 1)/n$ for $n = 1, 2, \ldots,$ since $A_{i,n}$ is a finite set for all $i < n$ and therefore, by finite additivity, $P(A_{i,n}) = 0$ for all $i < n$. Since $f_1, f_2, \ldots$ converges uniformly from below to f, $E(f, P) = 1$.

Now consider a sequence $g_1, g_2, \ldots$ that converges from below to g, but

not uniformly. In particular let $B_{1,n} = [0, 1/n] \cup ((n-1)/n, 1)$ and $B_{i,n} = ((i-1)/n, i/n] = A_{i,n}$ for $i = 2, \ldots, n-1$, and define g_n by

$$g_n(x) = \inf B_{i,n} \qquad \text{for all } x \in B_{i,n}, \qquad i = 1, \ldots, n-1.$$

Conditions 1 and 2 of Definition 10.11 hold for $g_1, g_2, \ldots$. But

1. $\sup \{E(g_n, P): n = 1, 2, \ldots\} \neq E(f, P)$ since $E(g_n, P) \equiv 0$;
2. Uniform convergence fails since for each n there are values of x for which $f(x) - g_n(x)$ is arbitrarily close to 1;
3. (10.3) of Lemma 10.2 fails since, with $\mathcal{B} = \{\{x: 0 \leq u(x) < c\}: 0 \leq c < 1\}$, $P(\bigcup_{\mathcal{B}} A) = P(X) = 1$ and $\sup \{P(A). A \in \mathcal{B}\} = 0$.

10.4 PREFERENCE AXIOMS AND BOUNDED UTILITIES

Because a number of conditions will be used in the theorems that follow we shall first summarize most of these conditions. In all cases, $\mathcal{A}$ is a Boolean algebra for X and $\mathcal{P}$ is a set of probability measures on $\mathcal{A}$. No notational distinction will be made between $x \in X$ and the one-point measure that assigns probability 1 to x. With $<$ defined on $\mathcal{P}$, $x < y \Leftrightarrow P < Q$ when $P(x) = Q(y) = 1$. Similar meanings hold for $x < P$, $x \lesssim P$, and so forth. As usual $P \lesssim Q \Leftrightarrow (P < Q$ or $P \sim Q)$, with $P \sim Q \Leftrightarrow$ (not $P < Q$, not $Q < P$).

We list first some primarily "structural" conditions.

S1. $\{x\} \in \mathcal{A}$ *for every* $x \in X$.
S2. $\{x: x \in X, x < y\} \in \mathcal{A}$ *and* $\{x: x \in X, y < x\} \in \mathcal{A}$ *for every* $y \in X$.
S3. $\mathcal{P}$ *contains every one-point probability measure.*
S4. $\mathcal{P}$ *is closed under countable convex combinations* (Definition 10.4).
S5. $\mathcal{P}$ *is closed under the formation of conditional probabilities* (Definition 10.8).

Conditions *S*1 and *S*3 enable us to define a utility function on X, and *S*2, which looks very much like some topological axioms of former chapters (see Theorems 3.5 and 5.5), guarantees that u on X is $\mathcal{A}$-measurable. In the present context $\mathcal{A}$ could be the set of all subsets of X (i.e., the discrete topology for X) and no problems would result. At worst we might have to deny countable additivity. On the other hand, the use of the discrete topology, which in general implies that $(X, \mathcal{T})$ is not connected, would have disastrous effects on former theory.

The following preference axioms (in addition to *S*2) include the three conditions of Chapter 8 along with three versions of a kind of dominance axiom. It is to be understood that these conditions apply to all $P, Q, R \in \mathcal{P}$, $A \in \mathcal{A}$, and $y, z \in X$.

*A*1. $<$ *on* $\mathcal{P}$ *is a weak order.*

*A*2. $(P < Q, 0 < \alpha < 1) \Rightarrow \alpha P + (1-\alpha)R < \alpha Q + (1-\alpha)R.$

*A*3. $(P < Q, Q < R) \Rightarrow \alpha P + (1-\alpha)R < Q$ *and* $Q < \beta P + (1-\beta)R$ *for some* $\alpha, \beta \in (0, 1)$.

*A*4*a*. $(P(A) = 1, Q < x$ *for all* $x \in A) \Rightarrow Q \leqslant P$. $(P(A) = 1, x < R$ *for all* $x \in A) \Rightarrow P \leqslant R$.

*A*4*b*. $(P(A) = 1, y \leqslant x$ *for all* $x \in A) \Rightarrow y \leqslant P$. $(P(A) = 1, x \leqslant z$ *for all* $x \in A) \Rightarrow P \leqslant z$.

*A*4*c*. $(P(A) = 1, y < x$ *for all* $x \in A) \Rightarrow y \leqslant P$. $(P(A) = 1, x < z$ *for all* $x \in A) \Rightarrow P \leqslant z$.

The final three conditions are weak versions of the following translation of Savage's *P*7 (1954, p. 77): $(P(A) = 1, Q \leqslant x$ for all $x \in A) \Rightarrow Q \leqslant P$, and $(P(A) = 1, x \leqslant R$ for all $x \in A) \Rightarrow P \leqslant R$. Axiom *A*4*a* is weaker (assumes less) than this since it replaces $\leqslant$ by $<$ in the hypotheses. *A*4*c* is weaker than *A*4*b* for the same reason. *A*4*b* is weaker than the translation of *P*7 since it deals only with one-point measures in part. Under *S*3, (*P*7 translation) $\Rightarrow$ (*A*4*a*, *A*4*b*, *A*4*c*), *A*4*a* $\Rightarrow$ *A*4*c*, and *A*4*b* $\Rightarrow$ *A*4*c*. Axiom *A*4*b* does not generally imply *A*4*a*, as can be seen from the proof of Theorem 10.2 in the next section. However, under the other hypotheses given above (*S*1–*A*3), *A*4*b* $\Rightarrow$ *A*4*a* when every $P \in \mathcal{P}$ is countably additive: this follows easily from Theorem 10.3. Under conditions *S*1–*A*3, *A*4*a* $\Rightarrow$ *A*4*b*.

In general, the dominance or sure-thing conditions *A*4*a*, *A*4*b*, and *A*4*c* seem reasonable, although *A*4*b* might be liable to criticism in the case where indifference is not transitive.

Bounded Consequence Utilities

The first result based on the new dominance conditions uses the weakest one of *A*4*a*, *A*4*b*, and *A*4*c*. We know from Theorem 8.4 that (10.14) and (10.15) follow from *A*1, *A*2, *A*3, and *S*4.

LEMMA 10.5. *Suppose that there is a real-valued function u on* $\mathcal{P}$ *for which*

$$P < Q \Leftrightarrow u(P) < u(Q), \qquad \text{for all } P, Q \in \mathcal{P}, \tag{10.14}$$

$$u(\alpha P + (1-\alpha)Q) = \alpha u(P) + (1-\alpha)u(Q), \qquad \text{for all } (\alpha, P, Q) \in [0, 1] \times \mathcal{P}^2, \tag{10.15}$$

*and suppose that S*1, *S*3, *S*4, *and A*4*c hold. Then, with* $u(x) = u(P)$ *when* $P(x) = 1$, *u on X is bounded.*

Proof. Under the hypotheses, suppose u on X is unbounded above. Then there are $x_1, x_2, \ldots$ such that $u(x_i) \geq 2^i$ for $i = 1, 2, \ldots$. By *S*4,

$\sum_{i=1}^{\infty} 2^{-i}x_{n+i} \in \mathcal{P}$ for $n = 0, 1, 2, \ldots$. By the easy extension of (10.15)

$$u\left(\sum_{i=1}^{\infty} 2^{-i}x_i\right) = \sum_{i=1}^{n} 2^{-i}u(x_i) + 2^{-n}u\left(\sum_{i=1}^{\infty} 2^{-i}x_{n+i}\right)$$

so that, since $u(x_i) \geq 2^i$,

$$u\left(\sum_{i=1}^{\infty} 2^{-i}x_i\right) \geq n + 2^{-n}u\left(\sum_{i=1}^{\infty} 2^{-i}x_{n+i}\right).$$

Since $y < x_i$ for all i greater than some m and for some $y \in X$, $A4c$ yields $y \preccurlyeq \sum_{i=1}^{\infty} 2^{-i}x_{n+i}$ for every $n \geq m$. Therefore, by (10.14),

$$u\left(\sum_{i=1}^{\infty} 2^{-i}x_i\right) \geq n + 2^{-n}u(y) \qquad \text{for } n = m+1, m+2, \ldots .$$

But this is false since $u(\sum_{i=1}^{\infty} 2^{-i}x_i)$ is a real number. Boundedness below is established by a symmetric contradiction. ◆

10.5 THEOREMS

By Theorem 8.4, the hypotheses of each theorem in this section imply the existence of a real-valued function u on $\mathcal{P}$ that satisfies (10.14) and (10.15) and is unique up to a positive linear transformation. As shown by Lemma 10.5, u on X is bounded. The question then is whether $u(P) = E(u, P)$ for all $P \in \mathcal{P}$, which is true if and only if there is a real-valued function u on X such that

$$P < Q \Leftrightarrow E(u, P) < E(u, Q), \qquad \text{for all } P, Q \in \mathcal{P}. \tag{10.16}$$

In the following theorems u is presumed to satisfy (10.14) and (10.15). These theorems show the weakest one of $A4a$, $A4b$, and $A4c$ that will yield (10.16) for various $\mathcal{P}$ sets. The $\not\Rightarrow$ means "do not imply for all possible cases." $H = \{S1, S2, S3, S4, S5, A1, A2, A3\}$.

THEOREM 10.1. $(H, A4a) \Rightarrow (10.16)$.

THEOREM 10.2. $(H, A4b) \not\Rightarrow (10.16)$.

THEOREM 10.3. (H, $A4b$, *every P is countably additive*) $\Rightarrow$ (10.16).

THEOREM 10.4. (H, $A4c$, *every P is countably additive*, $x < y$ *for some* $x, y \in X$) $\not\Rightarrow$ (10.16).

THEOREM 10.5. (H, $A4c$, *every P is discrete*, $x < y$ *for some* $x, y \in X$) $\Rightarrow$ (10.16).

THEOREM 10.6. *(H, $A4c$, every P is discrete)* $\not\Rightarrow$ (10.16).

The three "positive" theorems, Theorems 10.1, 10.3, and 10.5 are proved in the next section. The proofs of the three "negative" theorems are given in this section with specific cases where the hypotheses hold and (10.16) fails. These proofs illustrate some of the differences between measures that are not countably additive and those that are, and between countably additive measures that are not discrete and those that are.

Proof of Theorem 10.2. Let $X = \{0, 1, 2, \ldots\}$ with $u(x) = x/(1 + x)$ for all $x \in X$. Let $\mathcal{P}$ be the set of all probability measures on the set of all subsets of X and define u on $\mathcal{P}$ by

$$u(P) = E(u, P) + \inf \{P(u(x) \geq 1 - \epsilon) : 0 < \epsilon \leq 1\}.$$

The expression $P(u(x) \geq 1 - \epsilon)$ is a common shortening of $P(\{x : u(x) \geq 1 - \epsilon\})$. Define $<$ on $\mathcal{P}$ by $P < Q \Leftrightarrow u(P) < u(Q)$ so that (10.14) holds. By Exercises 6, 7, 8, and 18, (10.15) holds since

$$\begin{aligned}
u(\alpha P &+ (1 - \alpha)Q) \\
&= E(u, \alpha P + (1 - \alpha)Q) \\
&\quad + \inf\{\alpha P(u(x) \geq 1 - \epsilon) + (1 - \alpha)Q(u(x) \geq 1 - \epsilon) : 0 < \epsilon \leq 1\} \\
&= \alpha E(u, P) + (1 - \alpha)E(u, Q) + \alpha \inf\{P(u(x) \geq 1 - \epsilon) : 0 < \epsilon \leq 1\} \\
&\quad + (1 - \alpha) \inf\{Q(u(x) \geq 1 - \epsilon) : 0 < \epsilon \leq 1\} \\
&= \alpha u(P) + (1 - \alpha)u(Q).
\end{aligned}$$

H then follows from Theorem 8.4, and $A4b$ holds: if $P(A) = 1$ and $y \preccurlyeq x$ for all $x \in A$ then $u(y) \leq u(P)$ since $u(y) \leq E(u, P)$; if $P(A) = 1$ and $x \preccurlyeq z$ for all $x \in A$ then $u(P) \leq u(z)$ since $u(z) < 1 - \epsilon$ for some $\epsilon > 0$ and therefore $\inf\{P(u(x) \geq 1 - \epsilon) : 0 < \epsilon \leq 1\} = 0$.

Let P be diffuse with $P(x) = 0$ for all $x \in X$. Then $u(P) = 1 + 1 = 2$ since $\inf\{P(u(x) \geq 1 - \epsilon) : 0 < \epsilon \leq 1\} = 1$. Hence $u(P) \neq E(u, P)$. ◆

Proof of Theorem 10.4. Let $X = [0, 1]$, let $\mathcal{A}$ be the set of all Borel sets in $[0, 1]$. Take $\mathcal{P}$ as the set of countably-additive measures on $\mathcal{A}$. Set $u(x) = -1$ if $x < \frac{1}{2}$ and $u(x) = 1$ if $x \geq \frac{1}{2}$, and let

$$u(P) = \sum_X u(x)P(x), \qquad \text{for all } P \in \mathcal{P}. \tag{10.17}$$

u on $\mathcal{P}$ is well defined since $P(x) > 0$ for no more than a countable number of $x \in X$. Define $P < Q \Leftrightarrow u(P) < u(Q)$. Then (10.15) follows easily from

(10.17). The conditions in H hold and $A4c$ holds since

1. $(P(A) = 1, y < x \text{ for all } x \in A) \Rightarrow A \subseteq [\frac{1}{2}, 1], y \in [0, \frac{1}{2})$, and therefore $-1 = u(y) < 0 \leq u(P)$, and
2. $(P(A) = 1, x < z \text{ for all } x \in A) \Rightarrow A \subseteq [0, \frac{1}{2}), z \in [\frac{1}{2}, 1]$, and therefore $u(P) \leq 0 < u(z) = 1$.

But with Q the uniform measure on $[\frac{1}{2}, 1]$, $0 = u(Q) \neq E(u, Q) = 1$. ◆

Proof of Theorem 10.6. Let $X = \{1, 2, \ldots\}$, let $\mathcal{A}$ be the set of all subsets of X and let $\mathcal{P}$ be the set of all discrete probability measures on $\mathcal{A}$. Let $\mathcal{G}$ be the set of subsets of $\mathcal{P}$ defined by

$\mathcal{G} = \{\mathcal{S}: \mathcal{S} \subseteq \mathcal{P}$; if $P_1, \ldots, P_n, Q_1, \ldots, Q_m$ are all different measures in $\mathcal{S}$ and if $\alpha_i \geq 0, \beta_i \geq 0$ and $\sum_1^n \alpha_i = \sum_1^m \beta_j = 1$ then $\sum_1^n \alpha_i P_i \neq \sum_1^m \beta_j Q_j$; $\mathcal{S}$ contains all one-point measures$\}$.

A simple measure is in $\mathcal{S} \in \mathcal{G}$ only if it is a one-point measure. The measures in any $\mathcal{S} \in \mathcal{G}$ are independent with respect to finite convex combinations. A maximal independent subset is an $\mathcal{S}^* \in \mathcal{G}$ such that $\mathcal{S}^* \subset \mathcal{S}$ for no $\mathcal{S} \in \mathcal{G}$ and, if $P \in \mathcal{P}$ and $P \notin \mathcal{S}^*$ then there are positive numbers $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m$ with $\sum \alpha_i = \sum \beta_j = 1$ and distinct measures $P_2, \ldots, P_n, Q_1, \ldots, Q_m \in \mathcal{S}^*$ such that

$$\alpha_1 P + \sum_{i=2}^{n} \alpha_i P_i = \sum_{j=1}^{m} \beta_j Q_j \qquad (n \geq 1, m \geq 1). \tag{10.18}$$

Using Zorn's Lemma (Section 2.3) it is easily shown that $\mathcal{G}$ has a maximal element $\mathcal{S}^*$. It can be shown also, but is tedious algebraically to do so, that each $P \notin \mathcal{S}^*$ has an essentially unique representation in the form of (10.18).

If u is defined on the measures in $\mathcal{S}^*$, its *linear extension* to all of $\mathcal{P}$ is defined from (10.18) thus:

$$u(P) = \left[\sum_{j=1}^{m} \beta_j u(Q_j) - \sum_{i=2}^{n} \alpha_i u(P_i)\right] \Big/ \alpha_1.$$

To establish Theorem 10.6 define $u(x) = 0$ for all $x \in X$ and let $u(P) = 1$ for every $P \in \mathcal{S}^*$ that is not simple. Let u on $\mathcal{S}^*$ be extended linearly by (10.18) to all of $\mathcal{P}$ and define $P < Q \Leftrightarrow u(P) < u(Q)$. Then H is seen to hold and $A4c$ holds for the simple reason that $x < y$ for no $x, y \in X$. Hence the hypotheses of Theorem 10.6 hold. But (10.16) is clearly false. ◆

A variation on this example shows that $A4c$ cannot be deleted from the hypotheses of Theorem 10.5. Take $u(x) = x$ for each $x \in \{1, 2, \ldots\}$ and $u(P) = 0$ for each nonsimple $P \in \mathcal{S}^*$ and extend u linearly by (10.18) to all of $\mathcal{P}$. Define $P < Q \Leftrightarrow u(P) < u(Q)$. With $\gamma_1 R + \sum_{i=2}^{r} \gamma_i R_i = \sum_{j=1}^{s} \delta_j S_j$ along with P as in (10.18) we get $u(\alpha P + (1 - \alpha)R) = \alpha u(P) + (1 - \alpha)u(R)$ so

that (10.15) holds. Moreover, $x \prec y$ for some $x, y \in X$. But u on X is unbounded and therefore, by Lemma 10.5, *A4c* must be false. Clearly, (10.16) fails, for otherwise we could construct a P with infinite expected utility.

10.6 PROOFS OF THEOREMS 10.1, 10.3, AND 10.5

For Theorem 10.1 let H and *A4a* hold, let u on $\mathcal{P}$ satisfy (10.14) and (10.15), and define u on X as in Lemma 10.5. u on X is bounded. We note first that

$$P(A) = 1 \Rightarrow \inf\{u(x): x \in A\} \le u(P) \le \sup\{u(x): x \in A\}. \quad (10.19)$$

Let $c = \inf$ and $d = \sup$ in (10.19). To the contrary of (10.19), suppose that $d < u(P)$. Then, for any $x \in A$, (10.15) implies that there is a convex combination $R = \alpha P + (1 - \alpha)x$ such that $d < u(R) < u(P)$. Therefore, by (10.14), $x \prec R$ for all $x \in A$ and hence $P \precsim R$ by *A4a*. But this contradicts $u(R) < u(P)$. Hence $d < u(P)$ is false. $u(P) < c$ is seen to be false on using the other half of *A4a*. Hence (10.19) holds.

Let $a = \inf\{u(x): x \in X\}$ and $b = \sup\{u(x): x \in X\}$, and let $A_{i,n}$ be defined by (10.9). For $P \in \mathcal{P}$ let $n^* = \{i: i \in \{1, \ldots, n\}, P(A_{i,n}) > 0\}$. Then, by H and (10.7), $P = \sum_{n^*} P_{A_{i,n}} P(A_{i,n})$, so that $u(P) = \sum_{n^*} u(P_{A_{i,n}}) P(A_{i,n})$ by (10.15). Hence, by (10.9) and (10.19),

$$\sum_{n^*} [a + (i - 1)(b - a)/n] P(A_{i,n}) \le u(P) \le \sum_{n^*} [a + i(b - a)/n] P(A_{i,n}). \tag{10.20}$$

Since u is bounded and $\mathcal{A}$-measurable and $f_1, f_2, \ldots$ define by (10.10) converges uniformly from below to u, Definition 10.12 gives

$$E(u, P) = \sup\left\{\sum_{n^*} [a + (i - 1)(b - a)/n] P(A_{i,n}): n = 1, 2, \ldots\right\}.$$

Since the difference between the two sums in (10.20) equals $(b - a)/n$, which goes to 0 as n gets large, $u(P) = E(u, P)$. ◆

Proof of Theorem 10.3. Let H and *A4b* hold and assume that every $P \in \mathcal{P}$ is countably additive. With u as in the preceding proof, we need to verify (10.19). Then $u(P) = E(u, P)$ follows from the second half of the proof of Theorem 10.1.

Let $P(A) = 1$, $c = \inf\{u(x): x \in A\}$, and $d = \sup\{u(x): x \in A\}$. If $\{u(x): x \in A\} = \{c, d\}$, (10.19) follows from *A4b* and (10.14). Henceforth, assume that $c < u(w) < d$ for a fixed $w \in A$, and let

$$\begin{aligned} A_w &= \{x: x \in A, x \prec w\}, & \mathcal{P}_w &= \{Q: Q \in \mathcal{P}, Q(A_w) = 1\} \\ A^w &= \{x: x \in A, w \precsim x\}, & \mathcal{P}^w &= \{Q: Q \in \mathcal{P}, Q(A^w) = 1\}. \end{aligned} \tag{10.21}$$

$A = A_w \cup A^w$ with $A_w \ne \varnothing$ and $A^w \ne \varnothing$. Let $B = \{x: x \in X, x \prec w\}$ so that $B \in \mathcal{A}$ by *S2*. Then $A_w \in \mathcal{A}$ since $A_w = A \cap B = [A^c \cup B^c]^c$.

Similarly, $A^w \in \mathcal{A}$. Then, by (10.7) and *S*5, P equals a convex combination of a measure in $\mathcal{P}_w$ and a measure in $\mathcal{P}^w$. It follows from (10.15) that (10.19) holds if it holds for every measure in $\mathcal{P}_w \cup \mathcal{P}^w$.

To verify that $Q \in \mathcal{P}^w \Rightarrow c \leq u(Q) \leq d$, we note first that

$$c \leq u(Q) \qquad \text{for every } Q \in \mathcal{P}^w \tag{10.22}$$

follows from $c < u(w)$, *A4b*, (10.14), and (10.21). It follows from (10.22) and an analysis like that used in the proof of Lemma 10.5 that u on $\mathcal{P}^w$ is bounded above. Thus, let M be such that

$$c \leq u(Q) \leq M \qquad \text{for all } Q \in \mathcal{P}^w. \tag{10.23}$$

If $u(x) = d$ for some $x \in A^w$ then $u(Q) \leq d$ for all $Q \in \mathcal{P}^w$ by *A4b* and (10.14) so that $c \leq u(Q) \leq d$ for this case. Alternatively, suppose that $u(x) < d$ for all $x \in A^w$ and with $\epsilon > 0$ let

$$A(\epsilon) = \{x : x \in A^w, u(x) < d - \epsilon\}$$

$$B(\epsilon) = \{x : x \in A^w, d - \epsilon \leq u(x)\}.$$

Then $A(\epsilon) \cup B(\epsilon) = A^w$ and $\{A(\epsilon) : \epsilon > 0\}$ is weak ordered by $\subset$ so that for any $Q \in \mathcal{P}^w$ it follows from (10.3) of Lemma 10.2 that

$$\sup \{Q(A(\epsilon)) : \epsilon > 0\} = 1. \tag{10.24}$$

If $Q(A(\epsilon)) = 1$ for some $\epsilon > 0$ then $u(Q) < d$ by (10.14) and *A4b*. On the other hand, if $Q(A(\epsilon)) < 1$ for all $\epsilon > 0$ then, with ϵ small and $Q_{A(\epsilon)}$, $Q_{B(\epsilon)}$ respectively the conditional measure of Q given $A(\epsilon)$, $B(\epsilon)$, it follows from (10.15) and (10.7) that

$$u(Q) = Q(A(\epsilon))u(Q_{A(\epsilon)}) + Q(B(\epsilon))u(Q_{B(\epsilon)}).$$

Hence, by (10.23) and $Q_{A(\epsilon)}(A(\epsilon)) = 1$, $u(Q) < Q(A(\epsilon))\, d + [1 - Q(A(\epsilon))]M$ for all small $\epsilon > 0$. $u(Q) \leq d$ then follows from (10.24). Hence $Q \in \mathcal{P}^w \Rightarrow c \leq u(Q) \leq d$. By a symmetric proof, $Q \in \mathcal{P}_w \Rightarrow c \leq u(Q) \leq d$. ◆

Proof of Theorem 10.5. Let the hypotheses of Theorem 10.5 hold. Since every P is assumed to be discrete, $\mathcal{A}$ is a σ-algebra. With $a = \inf \{u(x) : x \in X\}$ and $b = \sup \{u(x) : x \in X\}$, $a < b$ since $x \prec y$ for some $x, y \in X$. We shall prove first that u on $\mathcal{P}$ is bounded.

If $a < u(w) < b$ for some $w \in X$, boundedness of u on $\mathcal{P}$ follows from an analysis like that using (10.22) in the preceding proof. Henceforth in this paragraph assume that $\{u(x) : x \in X\} = \{a, b\}$ and let

$$\mathcal{P}_a = \{P : P \in \mathcal{P}, P(u(x) = a) = 1\}$$

$$\mathcal{P}_b = \{P : P \in \mathcal{P}, P(u(x) = b) = 1\},$$

so that every $P \in \mathcal{P}$ is a convex combination of one measure from each of $\mathcal{P}_a$ and $\mathcal{P}_b$. (10.15) says that u on $\mathcal{P}$ is bounded if u is bounded on $\mathcal{P}_a \cup \mathcal{P}_b$. For $\mathcal{P}_b$, an analysis like that using (10.22) applies: $a \leq u(P)$ for all $P \in \mathcal{P}_b$ by $A4c$ and (10.14). A symmetric analysis shows that u on $\mathcal{P}_a$ is bounded.

If P is simple, $u(P) = E(u, P)$ follows from (10.15). If P is not simple and $A = \{x : P(x) > 0\}$, Lemma 10.3 gives $\sum_A P(x) = 1$. With the elements in A enumerated as $x_1, x_2, \ldots, P = \sum_{i=1}^{\infty} P(x_i)x_i$. Hence, by the finite extension of (10.15),

$$u(P) = \sum_1^n P(x_i)u(x_i) + \left[\sum_{n+1}^{\infty} P(x_i)\right] u\left(\sum_{n+1}^{\infty} P(x_i)\left[\sum_{n+1}^{\infty} P(x_i)\right]^{-1} x_i\right) \quad (10.25)$$

for $n = 1, 2, \ldots$. And by Exercise 20a,

$$E(u, P) = \sum_1^n P(x_i)E(u, x_i) + \left[\sum_{n+1}^{\infty} P(x_i)\right] E\left(u, \sum_{n+1}^{\infty} P(x_i)\left[\sum_{n+1}^{\infty} P(x_i)\right]^{-1} x_i\right) \quad (10.26)$$

for $n = 1, 2, \ldots$. Since $E(u, x_i) = u(x_i)$, it follows from (10.25) and (10.26) that

$$u(P) = E(u, P) + \left[\sum_{n+1}^{\infty} P(x_i)\right]\left[u\left(\sum_{n+1}^{\infty} \cdots x_i\right) - E\left(u, \sum_{n+1}^{\infty} \cdots x_i\right)\right]. \quad (10.27)$$

Since u on $\mathcal{P}$ is bounded, since $E(u, P)$ on $\mathcal{P}$ is bounded when u on X is bounded, and since $\sum_{n+1}^{\infty} P(x_i)$ approaches 0 as n gets large, the second term on the right of (10.27) approaches 0 as n gets large and therefore must equal zero for all n. Hence $u(P) = E(u, P)$. ◆

10.7 SUMMARY

The weak-order expected-utility result, $P < Q \Leftrightarrow E(u, P) < E(u, Q)$, holds for sets of probability measures that include nonsimple measures when appropriate dominance axioms are used. The basic idea of such axioms is that if a measure P is preferred to every consequence in a set to which a measure Q assigns probability 1, then Q shall not be preferred to P; and if every consequence in a set to which Q assigns probability 1 is preferred to P then P shall not be preferred to Q. This is condition $A4a$ of Section 10.4. If all probability measures under consideration are countably additive then a "weaker" form of dominance axiom will yield the expected-utility result in conjunction with the preference conditions of Chapter 8 and several structural conditions on the set of measures and the Boolean algebra on which they are defined.

INDEX TO EXERCISES

1. Denumerable sums. 2–3. Boolean algebras. 4. Countable unions. 5. σ-Algebra. 6–11. Infs and sups. 12–15. Countable additivity. 16. Uniform convergence from above. 17. Expectations of sums. 18–19. Expectations with convex combinations of measures. 20. Conditional expectations. 21. Expectations are sums. 22–24. Dominance and expectations. 25. *S*2. 26. Failure of *A*4*a*. 27–28. Proof of Theorem 10.6. 29. Blackwell-Girshick Theorem.

Exercises

1. Prove that $\sum_{i=1}^{n} 2^{-i} = 1 - 2^{-n}$ by noting that $2(\sum_1^n 2^{-i}) - \sum_1^n 2^{-i} = 1 - 2^{-n}$. Also show that $0 < p < 1$ implies $\sum_{n=1}^{\infty} p(1-p)^{n-1} = 1$.

2. Show that $\mathcal{A}$ is a Boolean algebra if and only if $\mathcal{A}$ is a nonempty set of subsets of X satisfying (2) and (3) of Definition 10.1.

3. Let $\mathcal{M}$ be the set of all subsets of $\{1, 2, \ldots\}$ that are either finite or contain all but a finite number of positive integers. Show that $\mathcal{M}$ is the Boolean algebra generated by $\{\{1\}, \{2\}, \ldots\}$.

4. Specify $\bigcup_{i=1}^{\infty} A_i = \{x : x \in A_i \text{ for some } i\}$ when (*a*) $A_i = \varnothing$, (*b*) $A_i = \{-i, i\}$, (*c*) $A_i = (1/(1+i), 1/i) \subseteq \mathrm{Re}$, (*d*) $A_i = [1/i, 2 - 1/i] \subseteq \mathrm{Re}$.

5. Describe the σ-algebra generated by $\{\{x\} : x \in \mathrm{Re}\}$.

6. Let R be a bounded set of numbers. Prove:
a. $\sup R = -\inf\{r : -r \in R\}$,
b. $\sup\{\alpha r : r \in R\} = \alpha \sup R$ if $\alpha \geq 0$,
c. $\sup\{\alpha r : r \in R\} = \alpha \inf R$ if $\alpha \leq 0$,
d. $\inf\{\alpha r : r \in R\} = \alpha \inf R$ if $\alpha \geq 0$,
e. $\inf\{\alpha r : r \in R\} = \alpha \sup R$ if $\alpha \leq 0$.

7. With R and S bounded sets of numbers prove that $\sup\{r + s : r \in R, s \in S\} = \sup R + \sup S$. Then prove Lemma 10.1.

8. (*Continuation.*) Prove that $\sup\{\alpha_i + \beta_i : i = 1, 2, \ldots\} = \sup\{\alpha_i : i = 1, 2, \ldots\} + \sup\{\beta_i : i = 1, 2, \ldots\}$ if $\alpha_1, \alpha_2, \ldots$ and $\beta_1, \beta_2, \ldots$ are nondecreasing sequences of real numbers that are bounded above. Generalize this result to n nondecreasing, bounded sequences.

9. (*Continuation.*) Suppose $\alpha_j \geq 0$ for all j, $\sum_{j=1}^{n} \alpha_j < M$ for all positive integers n and some number M, and for each j $(j = 1, 2, \ldots)$ $\beta_{1j}, \beta_{2j}, \ldots$ is a bounded nondecreasing sequence of nonnegative numbers. Using (10.1) prove that

$$\sup\left\{\sum_{j=1}^{\infty} \alpha_j \beta_{ij} : i = 1, 2, \ldots\right\} = \sum_{j=1}^{\infty} \alpha_j [\sup\{\beta_{ij} : i = 1, 2, \ldots\}].$$

10. (*Continuation.*) With the α_j as in the preceding exercise, suppose that, for

each j, $\gamma_{1j}, \gamma_{2j}, \ldots$ is a nonincreasing sequence of nonnegative real numbers. Prove that

$$\inf\left\{\sum_{j=1}^{\infty} \alpha_j \gamma_{ij} : i = 1, 2, \ldots\right\} = \sum_{j=1}^{\infty} \alpha_j [\inf\{\gamma_{ij} : i = 1, 2, \ldots\}].$$

11. Let $P_1, P_2, \ldots$ be a sequence of probability measures on the set of all subsets of X, let $\alpha_i \geq 0$ for all i with $\sum_{i=1}^{\infty} \alpha_i = 1$, and for any probability measure P and real-valued function u on X define

$$\lim_{\epsilon \to 0} P(u(x) \geq r - \epsilon) = \inf\{P(\{x : u(x) \geq r - \epsilon_j\}) : j = 1, 2, \ldots\}$$

where $\epsilon_1 > \epsilon_2 > \cdots$ and $\inf\{\epsilon_j : j = 1, 2, \ldots\} = 0$. Use the result of the preceding exercise to prove that, for any real number r,

$$\lim_{\epsilon \to 0} \sum_{i=1}^{\infty} \alpha_i P_i(u(x) \geq r - \epsilon) = \sum_{i=1}^{\infty} \alpha_i \left[\lim_{\epsilon \to 0} P_i(u(x) \geq r - \epsilon)\right].$$

12. Let P be defined on $\mathcal{M}$ of Exercise 3 on the basis of $P(n) = 2^{-n}$ for $n = 1, 2, \ldots$. Prove that P is countably additive.

13. Use the conclusion of Exercise 9 to prove (P_i is a countably-additive probability measure on $\mathcal{A}$ for $i = 1, 2, \ldots, \alpha_i \geq 0$ and $\sum_{i=1}^{\infty} \alpha_i = 1$) $\Rightarrow \sum_{i=1}^{\infty} \alpha_i P_i$ is a countably-additive probability measure on $\mathcal{A}$.

14. Prove that if P on $\mathcal{A}$ is countably additive, if $\mathcal{B}$ is a countable subset of $\mathcal{A}$ weakly ordered by $\subset$, and if $\bigcap_{\mathcal{B}} A \in \mathcal{A}$, then $P(\bigcap_{\mathcal{B}} A) = \inf\{P(A) : A \in \mathcal{B}\}$.

Note: In Exercises 15 *through* 24, $\mathcal{A}$ *is a Boolean algebra on* X; $f, g, \ldots$ *are bounded and* $\mathcal{A}$*-measurable;* $P, Q, \ldots$ *are probability measures on* $\mathcal{A}$.

15. If P is countably additive show that $E(f, P)$ is unambiguously defined by (10.11) when $f_1, f_2, \ldots$ is a sequence of simple $\mathcal{A}$-measurable functions that satisfies conditions 1 and 2 of Definition 10.11.

16. A sequence $g_1, g_2, \ldots$ of simple $\mathcal{A}$-measurable functions converges uniformly from above to f if and only if, for all $x \in X$,

1. $g_1(x) \geq g_2(x) \geq \cdots$
2. $g(x) = \inf\{g_i(x) : i = 1, 2, \ldots\}$
3. $\epsilon > 0 \Rightarrow g_n(x) \leq g(x) + \epsilon$ for some n (and all x).

Prove that $\sup\{E(f_n, P) : n = 1, 2, \ldots\} = \inf\{E(g_n, P) : n = 1, 2, \ldots\}$ when $f_1, f_2, \ldots (g_1, g_2, \ldots)$ converges uniformly from below (above) to f.

17. With c a real number let $f + c$ be the function on X that takes the value $f(x) + c$ at $x \in X$, let cf be the function that takes the value $cf(x)$ at $x \in X$, and let $f + g$ have value $f(x) + g(x)$ at x. Prove

a. $E(f + c, P) = E(f, P) + c$,
b. $E(f + g, P) = E(f, P) + E(g, P)$,
c. $E(cf, P) = cE(f, P)$.

18. Show that if $\alpha \in [0, 1]$ then

$$E(f, \alpha P + (1 - \alpha)Q) = \alpha E(f, P) + (1 - \alpha)E(f, Q) \qquad (10.28)$$

and then generalize this to $E(f, \sum_{i=1}^{n} \alpha_i P_i) = \sum_{i=1}^{n} \alpha_i E(f, P_i)$.

19. (*Continuation.*) Supposing that $\alpha_i \geq 0$ for $i = 1, 2, \ldots, \sum_{i=1}^{\infty} \alpha_i$ is finite, and $b_1, b_2, \ldots$ is a bounded sequence of numbers, define $\sum_{i=1}^{\infty} \alpha_i b_i$ from (10.1) by $\sum_{i=1}^{\infty} \alpha_i b_i = \sum_{i=1}^{\infty} \alpha_i(b_i + c) - c\sum_{i=1}^{\infty} \alpha_i$ where c is such that $b_i + c \geq 0$ for all i. Show that $\sum_{i=1}^{\infty} \alpha_i b_i$ is well defined. Then use this definition along with Exercise 17*a* and Exercise 9 to prove that if $\alpha_i \geq 0$ for all i and $\sum_{i=1}^{\infty} \alpha_i = 1$ then

$$E\left(f, \sum_{i=1}^{\infty} \alpha_i P_i\right) = \sum_{i=1}^{\infty} \alpha_i E(f, P_i). \tag{10.29}$$

20. Use the results of the two preceding exercises along with (10.7) and the sentence following its derivation to show that, given $A \in \mathcal{A}$,

a. $[P(A) = 1, \{A_1, \ldots, A_n\}$ is an $\mathcal{A}$-partition of A, $I = \{i: P(A_i) > 0\}] \Rightarrow E(f, P) = \sum_I P(A_i)E(f, P_{A_i})$,

b. $[P(A) = 1, \{A_1, A_2, \ldots\}$ is a denumerable partition of A with $A_i \in \mathcal{A}$ for all i, $I = \{i: P(A_i) > 0\}$, P is countably additive$] \Rightarrow E(f, P) = \sum_I P(A_i)E(f, P_{A_i})$.

21. Suppose all x_i are different in each of b and c. Show that

a. $P(x) = 1 \Rightarrow E(f, P) = f(x)$.

b. $\sum_{i=1}^{n} P(x_i) = 1 \Rightarrow E(f, P) = \sum_{i=1}^{n} P(x_i)f(x_i)$.

c. $\sum_{i=1}^{\infty} P(x_i) = 1 \Rightarrow E(f, P) = \sum_{i=1}^{\infty} P(x_i)f(x_i)$.

22. Assume that $A \in \mathcal{A}$ and $P(A) = 1$. Prove that

a. $[f(x) \leq g(x)$ for all $x \in A] \Rightarrow E(f, P) \leq E(g, P)$,

b. $[f(x) \leq g(x)$ for all $x \in A$, $P(f(x) + \epsilon \leq g(x)) > 0$ for some $\epsilon > 0] \Rightarrow E(f, P) < E(g, P)$,

c. $[f(x) < g(x)$ for all $x \in A$, P countably additive$] \Rightarrow E(f, P) < E(g, P)$.

23. (*Continuation.*) Give an example where $P(f(x) < g(x)) = 1$ and not $E(f, P) < E(g, P)$. $P(f(x) < g(x)) = P(\{x: f(x) < g(x)\})$.

24. With u satisfying (10.16) let $A_y = \{x: x \in X, x \preccurlyeq y\}$. Prove that $[P(A_y) \leq Q(A_y)$ for all $y \in X] \Rightarrow E(u, Q) \leq E(u, P)$. Prove also that $[Q \neq P_i$ for $i = 1, 2, \ldots, n$, $\alpha_i \geq 0$ for all $i, \sum_1^n \alpha_i = 1, \sum_{i=1}^{n} \alpha_i P_i(A_y) \leq Q(A_y)$ for all $y \in X] \Rightarrow E(u, Q) \leq E(u, P_i)$ for some i.

25. Show that Definition 10.1, *S*2, and $\preccurlyeq$ on X connected imply $\{x: x \in X, y \preccurlyeq x \preccurlyeq z\} \in \mathcal{A}$ and $\{x: x \in X, y \prec x \prec z\} \in \mathcal{A}$.

26. Give specific examples of probability measures that demonstrate the failure of *A*4*a* in the proofs of Theorems 10.2 and 10.4.

27. Use Zorn's Lemma to prove that $\mathcal{S}$ in the proof of Theorem 10.6 has a maximal element $\mathcal{S}^*$.

28. Verify that the representation (10.18) for $(P \in \mathcal{P}, P \notin \mathcal{S}^*)$ in terms of measures in $\mathcal{S}^*$ is essentially unique.

29. Blackwell and Girshick (1954). Prove that if $\mathcal{P}$ is the set of all discrete probability measures on the set of all subsets of X and if *A*1 and *A*3 of Section 10.4 hold along with $[P_i, Q_i \in \mathcal{P}, P_i \preccurlyeq Q_i$ and $\alpha_i \geq 0$ for $i = 1, 2, \ldots; \sum_{i=1}^{\infty} \alpha_i = 1$; $P_i \prec Q_i$ for some i for which $\alpha_i > 0] \Rightarrow \sum_{i=1}^{\infty} \alpha_i P_i \prec \sum_{i=1}^{\infty} \alpha_i Q_i$, then there is a real-valued function u on X that satisfies (10.16).

Chapter 11

ADDITIVE EXPECTED UTILITY

This chapter combines the weak-order expected utility theory of Chapters 8 and 10 with the situation where the consequences in X are n-tuples as in Chapters 4, 5, and 7. The main focus of the chapter is conditions that, when $X \subseteq X_1 \times X_2 \times \cdots \times X_n$, imply the existence of real-valued functions $u_1, \ldots, u_n$ on $X_1, \ldots, X_n$ such that

$$P < Q \Leftrightarrow \sum_{i=1}^{n} E(u_i, P_i) < \sum_{i=1}^{n} E(u_i, Q_i), \qquad \text{for all } P, Q \in \mathcal{P}, \qquad (11.1)$$

where $\mathcal{P}$ is a set of probability measures on X and, for $P \in \mathcal{P}$, P_i is the marginal measure of P on X_i.

We shall examine (11.1) first for the case where $X = X_1 \times \cdots \times X_n$ and then for the more general case where $X \subseteq X_1 \times \cdots \times X_n$. Section 11.3 then examines the case where $X = X_1 \times \cdots \times X_n$ and (11.1) may fail but some form of additive interdependence applies such as $u(x_1, x_2, x_3) = u_1(x_1, x_2) + u_2(x_2, x_3)$. Finally, Section 11.4 looks at the homogeneous product set situation where $X = A^n$, as in Chapter 7.

11.1 ADDITIVE EXPECTED UTILITY WITH $X = \Pi X_i$

To simplify our examination of independence among factors in a multidimensional consequence set in the expected-utility context, this chapter assumes that probability measures for X are defined on the set of all subsets of X. A similar assumption applies for a measure defined for a factor set X_i.

Definition 11.1. Suppose P is a probability measure on $X \subseteq \prod_{i=1}^{n} X_i$. Then P_i, the *marginal measure* of P on X_i, is defined by

$$P_i(A_i) = P(\{x : x \in X, x_i \in A_i\}) \qquad \text{for all } A_i \subseteq X_i. \qquad (11.2)$$

In (11.2) x_i is the ith component of x. When $X = \prod_{i=1}^{n} X_i$, (11.2) becomes

$P_i(A_i) = P(X_1 \times \cdots \times X_{i-1} \times A_i \times X_{i+1} \times \cdots \times X_n)$. It is easily verified that P_i is a probability measure on X_i when P is a probability measure on X.

It is possible to have $(P_1, \ldots, P_n) = (Q_1, \ldots, Q_n)$ when $P \neq Q$. With $n = 2$ let P and Q be the simple even-chance gambles

$$P(\$5000, \$5000) = P(\$100000, \$100000) = .5$$

$$Q(\$5000, \$100000) = Q(\$100000, \$5000) = .5.$$

Then $(P_1, P_2) = (Q_1, Q_2)$ although $P \neq Q$. P gives an even chance for a two-year income stream of either (\$5000, \$5000) or (\$100000, \$100000). Q gives an even chance for an income stream of either (\$5000, \$100000) or (\$100000, \$5000). I suspect that many people would prefer Q to P. The condition for (11.1) requires that $P \sim Q$. This condition may seem more reasonable when the different factors in X are heterogeneous.

THEOREM 11.1. *Suppose that* $\mathcal{P}$ *is either the set of simple probability measures on* $X = \prod_{i=1}^n X_i$ *or a set of probability measures on* $X = \prod_{i=1}^n X_i$ *that satisfies S1 through S5 of Section* 10.4, *and suppose further that there is a real-valued function u on X such that, for all* $P, Q \in \mathcal{P}$, $P < Q \Leftrightarrow E(u, P) < E(u, Q)$. *Then there are real-valued functions* $u_1, \ldots, u_n$ *on* $X_1, \ldots, X_n$ *respectively that satisfy* (11.1) *and are unique up to similar positive linear transformations, if and only if* $P \sim Q$ *whenever P and Q are simple measures in* $\mathcal{P}$ *such that* $(P_1, \ldots, P_n) = (Q_1, \ldots, Q_n)$ *and* $P(x), Q(x) \in \{0, \frac{1}{2}, 1\}$ *for all* $x \in X$.

The very last condition here shows that (11.1) can be established on the basis of simple 50-50 gambles when $X = \prod_{i=1}^n X_i$. The necessity of the indifference condition for (11.1) is obvious. The sufficiency proof follows.

Proof. Fix $x^0 = (x_1^0, \ldots, x_n^0)$ in X, assign $u_1(x_1^0), \ldots, u_n(x_n^0)$ values that sum to $u(x^0)$, and define u_i on X_i by

$$u_i(x_i) = u(x_1^0, \ldots, x_{i-1}^0, x_i, x_{i+1}^0, \ldots, x_n^0) - \sum_{j \neq i} u_j(x_j^0). \tag{11.3}$$

The indifference condition between 50-50 gambles when $(P_1, \ldots, P_n) = (Q_1, \ldots, Q_n)$ leads directly to $u(x_1, \ldots, x_i, x_{i+1}^0, \ldots, x_n^0) + u(x_1^0, \ldots, x_i^0, x_{i+1}, x_{i+2}^0, \ldots, x_n^0) = u(x_1, \ldots, x_{i+1}, x_{i+2}^0, \ldots, x_n^0) + u(x^0)$ for $i = 1, \ldots, n - 1$. Summing this from $i = 1$ to $i = n - 1$, cancelling identical terms, and transposing $(n - 1)u(x^0)$ we get $\sum_{i=1}^n u(x_1^0, \ldots, x_{i-1}^0, x_i, x_{i+1}^0, \ldots, x_n^0) - (n - 1)u(x_1^0, \ldots, x_n^0) = u(x_1, \ldots, x_n)$, which on comparison with (11.3) shows that

$$u(x_1, \ldots, x_n) = \sum_{i=1}^n u_i(x_i), \qquad \text{for all } (x_1, \ldots, x_n) \in X. \tag{11.4}$$

If $\mathcal{T}$ satisfies the Section 10.4 conditions then u on X is bounded and hence u_i on X_i, defined by (11.3), is bounded. In any event, although u_i is defined on X_i, it is equivalent to a function u_i^* defined on X by $u_i^*(x) = u_i(x_i)$. Then, using (11.4) and Exercise 10.17*b*,

$$\begin{aligned} E(u, P) &= E(u_1^* + \cdots + u_n^*, P) \\ &= E(u_1^*, P) + \cdots + E(u_n^*, P) \\ &= E(u_1, P_1) + \cdots + E(u_n, P_n) \end{aligned}$$

which yields (11.1) in conjunction with $P < Q \Leftrightarrow E(u, P) < E(u, Q)$.

Finally, suppose that $v_1, \ldots, v_n$ on $X_1, \ldots, X_n$ satisfy (11.1) along with $u_1, \ldots, u_n$. Define u and v on the simple measures in $\mathcal{T}$ by $u(P) = \sum_i E(u_i, P_i)$ and $v(P) = \sum_i E(v_i, P_i)$. It is easily seen that $u(\alpha P + (1 - \alpha)Q) = \alpha u(P) + (1 - \alpha)u(Q)$ and similarly for v for simple measures $P, Q \in \mathcal{T}$. Hence, by Theorem 8.4, v is a positive linear transformation of u, say $v = au + b$, $a > 0$. We then have $\sum v_i(x_i) = \sum E(v_i, x_i) = v(x_1, \ldots, x_n) = au(x_1, \ldots, x_n) + b = a \sum u_i(x_i) + b$, from which it follows that $v_i(x_i) = au_i(x_i) + [b + a \sum_{j \neq i} u_j(x_j^0) - \sum_{j \neq i} v_j(x_j^0)] = au_i(x_i) + b_i$, for each i, where b_i is defined in context. ◆

11.2 ADDITIVE EXPECTATIONS WITH $X \subseteq \Pi X_i$

When $X \subseteq \prod_{i=1}^n X_i$, (11.1) does not generally follow from the 50-50 gambles version of the indifference condition. In general, we require the more general condition that $(P_1, \ldots, P_n) = (Q_1, \ldots, Q_n) \Rightarrow P \sim Q$. When X is finite, (11.1) follows from this and $P < Q \Leftrightarrow E(u, P) < E(u, Q)$ as is noted in Exercise 4. For X infinite, only the $n = 2$ case has been satisfactorily worked out and then only for simple probability measures. Therefore, this section examines only the $X \subseteq X_1 \times X_2$ case.

To show one difficulty that may arise in this case for nonsimple probability measures suppose $X = \{(0, 0), (1, 0), (1, 1), (2, 1), (2, 2), (3, 2), \ldots\}$ and let u on X, satisfying $P < Q \Leftrightarrow E(u, P) < E(u, Q)$ for all discrete measures P and Q on X, be such that $u(k, k) = 0$ for $k = 0, 1, 2, \ldots$ and $u(k + 1, k) = 1$ for $k = 0, 1, 2, \ldots$. Set $u_1(0) = u_2(0) = 0$ for (11.1). Then for (11.1) to hold for all one-point probability measures we must have $u_1(k) = k$ and $u_2(k) = -k$ for $k = 0, 1, 2, \ldots$ [when $u(x_1, x_2) = u_1(x_1) + u_2(x_2)$]. Define P by

$$P(2^k, 2^k) = 2^{-k} \qquad \text{for } k = 1, 2, \ldots$$

so that $E(u, P) = 0$. Then $E(u_1, P)$, if defined at all, is infinite: $E(u_1, P) = 2^1 \cdot 2^{-1} + 2^2 \cdot 2^{-2} + \cdots = 1 + 1 + \cdots = +\infty$. Likewise $E(u_2, P) = -2^1 \cdot 2^{-1} - 2^2 \cdot 2^{-2} - \cdots = -\infty$, so that $E(u_1, P_1) + E(u_2, P_2)$ is not meaningfully defined. Despite this, $E(u, Q) = E(u_1, Q_1) + E(u_2, Q_2)$ when Q is simple.

The Structure of $X \subseteq X_1 \times X_2$

Several special definitions that apply to this section only will be used in resolving the $X \subseteq X_1 \times X_2$ case for simple probability measures.

Definition 11.2. $(x_1, x_2)R(y_1, y_2)$ if and only if there is a finite sequence $(x_1, x_2), x^1, \ldots, x^N, (y_1, y_2)$ of elements in $X \subseteq X_1 \times X_2$ such that any two adjacent elements have at least one component in common.

R is easily seen to be an equivalence on X. We shall let $\mathfrak{D}$ be the set of equivalence classes of X under R. In the preceding example $\mathfrak{D} = \{X\}$. If $D, D^* \in \mathfrak{D}$ and $(x_1, x_2) \in D$, $(y_1, y_2) \in D^*$, and $D \neq D^*$ then it must be true that $x_1 \neq x_2$ and $y_1 \neq y_2$. Hence, given u on X, there will be a u_1, u_2 solution to

$$u(x_1, x_2) = u_1(x_1) + u_2(x_2) \qquad \text{for all } (x_1, x_2) \in X \tag{11.5}$$

as required for (11.1) if and only if there is a u_1, u_2 solution for each $D \in \mathfrak{D}$ considered separately. We therefore concentrate on an arbitrary $D \in \mathfrak{D}$. Proofs of our first three lemmas are left to the reader.

Definition 11.3. An *alternating sequence* in D is a finite sequence of two or more *distinct* elements in D such that

1. any two adjacent elements have one component in common,
2. no three consecutive elements in the sequence have the same first component or the same second component.

LEMMA 11.1. *If $x, y \in D$ and $x \neq y$ then there is an alternating sequence in D that begins with x and ends with y.*

Definition 11.4. A *cycle* in D is a subset of an *even* number of elements in D that can be positioned in an alternating sequence whose first and last elements have the same first component if the first and second elements have the same second component or whose first and last elements have the same second component if the first and second elements have the same first component.

LEMMA 11.2. *If D has no cycles then there is exactly one alternating sequence in D from x to y when $x, y \in D$ and $x \neq y$.*

LEMMA 11.3. *Suppose $x^1, \ldots, x^{2n}$ is an alternating sequence in D whose elements form a cycle. Suppose further that $\mathcal{P}$ is a mixture set of probability measures that includes the simple measures, that there is a real-valued function*

u on $\mathcal{P}$ that satisfies (8.5) *and* (8.6), *and that* $[P, Q \in \mathcal{P}, (P_1, P_2) = (Q_1, Q_2)] \Rightarrow P \sim Q$. *Then*

$$\sum_{i=1}^{n} u(x_1^{2i-1}, x_2^{2i-1}) = \sum_{i=1}^{n} u(x_1^{2i}, x_2^{2i}). \tag{11.6}$$

LEMMA 11.4. *If $D \in \mathfrak{D}$ then there is a $C \subseteq D$ such that*

1. *C includes no cycle,*
2. *xRy for each $x, y \in C$ with xRy established by a sequence all of whose elements are in C,*
3. $(x_1, x_2) \in D \Rightarrow (x_1, y_2) \in C$ *and* $(y_1, x_2) \in C$ *for some* $y_1 \in X_1$, $y_2 \in X_2$.

Proof. With $D \in \mathfrak{D}$ let

$$\mathcal{C} = \{C : C \subseteq D, C \text{ satisfies conditions 1 and 2 of Lemma 11.4}\}.$$

We shall prove that $\mathcal{C}$ has a maximal element that satisfies condition 3. Let $\mathcal{C}^*$ be a subset of $\mathcal{C}$ that is strictly ordered by $\subset$, and let $C^* = \bigcup_{\mathcal{C}^*} C$. $C^* \in \mathcal{C}$ since $C^* \subseteq D$ and

1. C^* includes no cycle, for if $\{x^1, \ldots, x^n\} \subseteq C^*$ is a cycle then with $x^i \in C_i$, $C_i \in \mathcal{C}^*$, the largest of these C_i will include $\{x^1, \ldots, x^n\}$ and this contradicts $C_i \in \mathcal{C}$;
2. xRy if $x, y \in C^*$, for $x, y \in C$ for some $C \in \mathcal{C}^*$. Thus, by Zorn's Lemma, there is a $B \in \mathcal{C}$ such that $B \subset C$ for no $C \in \mathcal{C}$. Suppose $(x_1, x_2) \in D$ and $(x_1, x_2) \notin B$. Then, since B is maximal, $B \cup \{(x_1, x_2)\}$ must include a cycle which, since B includes no cycles, contains (x_1, x_2). It follows from Definitions 11.3 and 11.4 that $(x_1, y_2) \in B$ and $(y_1, x_2) \in B$ for some $y_2 \in X_2$ and $y_1 \in X_1$. ◆

Additive Expected Utility with Simple Measures

The appropriate theorem for (11.1) with $X \subseteq X_1 \times X_2$ and simple probability measures follows.

THEOREM 11.2. *Suppose $\mathcal{P}$ is the set of simple probability measures on $X \subseteq X_1 \times X_2$ and there is a real-valued function u on X such that $P < Q \Leftrightarrow E(u, P) < E(u, Q)$, for all $P, Q \in \mathcal{P}$. Then there are real-valued functions u_1 on X_1 and u_2 on X_2 that satisfy* (11.1) *if and only if* $[P, Q \in \mathcal{P}, (P_1, P_2) = (Q_1, Q_2)] \Rightarrow P \sim Q$. *If v_1 on X_1 and v_2 on X_2 satisfy* (11.1) *along with u_1 and u_2 then there are numbers $a > 0$ and b and real-valued functions f_1 and f_2 on $\mathfrak{D}$ such that*

$$\begin{aligned} v_1(x_1) &= au_1(x_1) + f_1(D(x_1)) && \textit{for all } x_1 \in X_1 \\ v_2(x_2) &= au_2(x_2) + f_2(D(x_2)) && \textit{for all } x_2 \in X_2 \\ f_1(D) + f_2(D) &= b && \textit{for all } D \in \mathfrak{D} \end{aligned}$$

where $D(x_i) \in \mathfrak{D}$ contains an element whose i*th component is x_i.*

For completeness we should mention that each $x_i \in X_i$ is assumed to be the ith component of some $x \in X$, for $i = 1, 2$.

Proof. The sufficiency of the hypotheses for (11.1) is proved by showing that (11.5) holds. Two cases are considered for any $D \in \mathfrak{D}$.

Case 1: D has no cycles. Fix $x^0 \in D$, define $u_1(x_1^0)$ and $u_2(x_2^0)$ so that $u_1(x_1^0) + u_2(x_2^0) = u(x^0)$, and proceed term by term along alternating sequences beginning at x^0, defining u_1 and u_2 in the only way possible to satisfy (11.5). By Lemma 11.1, every x_1 and x_2 in elements in D has a $u_1(x_1)$ or $u_2(x_2)$ thus defined. Lemma 11.2 implies that the u_1 and u_2 values are unique, given $u_1(x_1^0)$ and $u_2(x_2^0)$.

Case 2: D has cycles. Let $C \subseteq D$ satisfy the three conditions of Lemma 11.4. The Case 1 proof then applies to C and gives $u(x_1, x_2) = u_1(x_1) + u_2(x_2)$ for all $x \in C$. Suppose $x \in D$, $x \notin C$. Then, by condition 3 of Lemma 11.4, we have $(x_1, y_2) \in C$ and $(y_1, x_2) \in C$ and, by Lemmas 11.1 and 11.2 there is a unique alternating sequence in C from (x_1, y_2) to (y_1, x_2). Hence $C \cup \{x\}$ has a cycle that must include x. An alternating sequence whose elements form such a cycle can be written as $(x_1, x_2), (x_1^2, x_2^2), \ldots, (x_1^{2n}, x_2^{2n})$. By Lemma 11.3, (11.6) holds with $(x_1^1, x_2^1) = (x_1, x_2)$. Applying $u = u_1 + u_2$ to the C terms in the cycle it follows from (11.6) after cancellation that $u(x_1, x_2) = u_1(x_1) + u_2(x_2)$. It follows that $u = u_1 + u_2$ holds on all of D.

For the last part of the theorem let u_1, u_2 and v_1, v_2 each satisfy (11.1). Using the approach in the final paragraph of the proof of Theorem 11.1 we get $v_1(x_1) + v_2(x_2) = au_1(x_1) + au_2(x_2) + b$ for all $x \in X$. For a given D let $x^0 \in D$. The Case 1 procedure for assigning u_1 and u_2 then leads to

$$v_1(x_1) = au_1(x_1) + au_2(x_2^0) + b - v_2(x_2^0)$$
$$v_2(x_2) = au_2(x_2) + au_1(x_1^0) + b - v_1(x_1^0)$$

for all $x \in D$. Letting $f_1(D) = au_2(x_2^0) + b - v_2(x_2^0)$ and $f_2(D) = au_1(x_1^0) + b - v_1(x_1^0)$, the desired equations follow. (A different x^0 must be chosen for each D since there are no x_1 or x_2 interconnections between different elements in $\mathfrak{D}$.) ◆

11.3 ADDITIVE, INTERDEPENDENT EXPECTATIONS FOR ΠX_i

Throughout this section we take $X = \prod_{i=1}^n X_i$ and let $\{I_1, \ldots, I_m\}$ be an arbitrary, but fixed, nonempty set of nonempty subsets of $\{1, 2, \ldots, n\}$. For Section 11.1, $I_j = \{j\}$ for $j = 1, \ldots, n$. Here we shall permit the I_j to overlap.

We shall let $\mathcal{P}$ be a set of probability measures on X and let $\mathcal{P}_j$ be a set of probability measures on $\prod_{i \in I_j} X_i$. With $P \in \mathcal{P}$, the *marginal measure* P_j of P on $\prod_{I_j} X_i$ is such that $P_j(A_j) = P(\{x : x \in X, x_i \in X_i \text{ for } i \notin I_j, x^j \in A_j\})$ for

every $A_j \subseteq \prod_{I_j} X_i$, where x^j is the *projection* of x onto I_j. For example, if $x = (x_1, x_2, x_3, x_4)$ and $I_j = \{1, 3\}$ then $x^j = (x_1, x_3)$.

THEOREM 11.3. *Suppose that the hypotheses in the first sentence of Theorem* 11.1 *hold. Then there are real-valued functions* $u_1, \ldots, u_m$ *on* $\prod_{I_1} X_i, \ldots, \prod_{I_m} X_i$ *respectively such that*

$$P < Q \Leftrightarrow \sum_{j=1}^{m} E(u_j, P_j) < \sum_{j=1}^{m} E(u_j, Q_j), \qquad \text{for all } P, Q \in \mathfrak{P}, \tag{11.7}$$

if and only if $[P, Q \in \mathfrak{P}, (P_1, \ldots, P_m) = (Q_1, \ldots, Q_m)] \Rightarrow P \sim Q$.

Admissible transformations for the u_j are discussed in Exercises 9c and 10. The proof of Theorem 11.3 will be carried out in two steps. First, we shall state and prove a lemma and then use this to prove the theorem. In the statement of the lemma we shall let $x^0 = (x_1^0, \ldots, x_n^0)$ in X be fixed and, for any $x \in X$ and $I \subseteq \{1, 2, \ldots, n\}$, let $x[I]$ be the n-tuple in X whose ith component is x_i if $i \in I$ and x_i^0 if $i \notin I$.

LEMMA 11.5. *Suppose* $\mathfrak{P}$ *contains every simple probability measure on* X, *and a real-valued function* u *on* X *satisfies* $P < Q \Leftrightarrow E(u, P) < E(u, Q)$ *for all simple measures. Suppose further that* $[P, Q \in \mathfrak{P}, (P_1, \ldots, P_m) = (Q_1, \ldots, Q_m)] \Rightarrow P \sim Q$. *Then, for all* $x \in X$,

$$u(x) = \sum_{j=1}^{m} (-1)^{j+1} \sum_{1 \le i_1 < i_2 < \cdots < i_j \le m} u\left(x\left[\bigcap_{s=1}^{j} I_{i_s}\right]\right). \tag{11.8}$$

For $m = 3$, (11.8) is $u(x) = u(x[I_1]) + u(x[I_2]) + u(x[I_3]) - \{u(x[I_1 \cap I_2]) + u(x[I_1 \cap I_3]) + u(x[I_2 \cap I_3])\} + u(x[I_1 \cap I_2 \cap I_3])$.

Proof of Lemma 11.5. To simplify notation let x be an arbitrary element in X and let $[I]^j$ be the projection of $x[I]$ onto I_j. (If $I = \{1, 3\}$ then $x[I] = (x_1, x_2^0, x_3, x_4^0, \ldots)$. Then if $I_j = \{1, 4\}$, $[I]^j = (x_1, x_4^0)$.) Because the only integers in I that are relevant in defining $[I]^j$ are those in I_j, $[I]^j = [I \cap I_j]^j$.

Let S and R on $\{1, \ldots, m\} \times \{1, \ldots, m\}$ be defined by

$$S(k,j) = \{(i_1, \ldots, i_k): 1 \le i_1 < \cdots < i_k \le m, j \in \{i_1, \ldots, i_k\}\}$$
$$R(k,j) = \{(i_1, \ldots, i_k): 1 \le i_1 < \cdots < i_k \le m, j \notin \{i_1, \ldots, i_k\}\}$$

so that $S(1,j) = \{j\}$, $R(m,j) = \varnothing$, and

$$S(k,j) \cup R(k,j) = \{(i_1, \ldots, i_k): 1 \le i_1 < \cdots < i_k \le m\}, \qquad j = 1, \ldots, m.$$

$S(k,j) \cup R(k,j)$ has $\binom{m}{k}$ elements.

Let P and Q be simple probability measures defined by

$$P = \alpha x + \sum_{k \in K_e} \sum_{S(k,j) \cup R(k,j)} \alpha x\left[\bigcap_{s=1}^{k} I_{i_s}\right]$$

$$Q = \sum_{k \in K_0} \sum_{S(k,j) \cup R(k,j)} \alpha x\left[\bigcap_{s=1}^{k} I_{i_s}\right]$$

where $K_e = \{i: 2 \le i \le m,\ i \text{ is } even\}$, $K_0 = \{i: 1 \le i \le m,\ i \text{ is } odd\}$, and $\alpha = 2^{-m+1}$. For the marginal measures on $\prod_{I_j} X_i$ our preceding observations and definitions yield

$$\begin{aligned}
P_j &= \alpha x^j + \sum_{K_e}\left(\sum_{S(k,j)} \alpha\left[\bigcap_{s=1}^{k} I_{i_s}\right]^j + \sum_{R(k,j)} \alpha\left[\bigcap_{s=1}^{k} I_{i_s}\right]^j\right) \\
&= \alpha x^j + \sum_{K_e}\left(\sum_{R(k-1,j)} \alpha\left[\bigcap_{s=1}^{k-1} I_{i_s}\right]^j + \sum_{R(k,j)} \alpha\left[\bigcap_{s=1}^{k} I_{i_s}\right]^j\right) \\
&= \alpha x^j + \sum_{k=1}^{m} \sum_{R(k,j)} \alpha\left[\bigcap_{s=1}^{k} I_{i_s}\right]^j \\
&= \alpha x^j + \sum_{R(1,j)} \alpha[I_{i_1}]^j + \sum_{\substack{k \in K_0 \\ k \ge 3}}\left(\sum_{R(k-1,j)} \alpha\left[\bigcap_{s=1}^{k-1} I_{i_s}\right]^j + \sum_{R(k,j)} \alpha\left[\bigcap_{s=1}^{k} I_{i_s}\right]^j\right) \\
&= \sum_{S(1,j)} \alpha[I_{i_1}]^j + \sum_{R(1,j)} \alpha[I_{i_1}]^j + \sum_{\substack{k \in K_0 \\ k \ge 3}}\left(\sum_{S(k,j)} \alpha\left[\bigcap_{s=1}^{k} I_{i_s}\right]^j\right. \\
&\qquad \left. + \sum_{R(k,j)} \alpha\left[\bigcap_{s=1}^{k} I_{i_s}\right]^j\right) \\
&= \sum_{K_0} \sum_{S(k,j) \cup R(k,j)} \alpha\left[\bigcap_{s=1}^{k} I_{i_s}\right]^j \\
&= Q_j.
\end{aligned}$$

Hence, by hypothesis, $P \sim Q$ and therefore $E(u, P) = E(u, Q)$, or

$$u(x) + \sum_{K_e} \sum_{S(k,j) \cup R(k,j)} u\left(x\left[\bigcap_{s=1}^{k} I_{i_s}\right]\right) = \sum_{K_0} \sum_{S(k,j) \cup R(k,j)} u\left(x\left[\bigcap_{s=1}^{k} I_{i_s}\right]\right)$$

which is (11.8). ◆

Proof of Theorem 11.3 (*Sufficiency*). To verify (11.7) under the stated hypotheses, including $(P_1, \ldots, P_m) = (Q_1, \ldots, Q_m) \Rightarrow P \sim Q$, we note first from Lemma 11.5 that (11.8) holds for u on X. With $x^0 \in X$ fixed as in the lemma we define u_j on $\prod_{I_j} X_i$ as follows:

$$u_j(x^j) = u(x[I_j]) + \sum_{k=1}^{j-1}(-1)^k \sum_{1 \le i_1 < \cdots < i_k < j} u\left(x\left[\bigcap_{s=1}^{k} I_{i_s} \cap I_j\right]\right).$$

u_j is well defined since $u_j(x^j) = u_j(y^j)$ if $x^j = y^j$. Moreover, if u on X is bounded (as in Chapter 10) then u_j is bounded. Summing over j:

$$\begin{aligned}
\sum_{j=1}^{m} u_j(x^j) &= \sum_{j=1}^{m} u(x[I_j]) + \sum_{j=1}^{m}\sum_{k=1}^{j-1}(-1)^k \sum_{1\le i_1<\cdots<i_k<j} u\left(x\left[\bigcap_{s=1}^{k} I_{i_s} \cap I_j\right]\right) \\
&= \sum_{j=1}^{m} u(x[I_j]) + \sum_{k=1}^{m-1}(-1)^k \sum_{j=k+1}^{m}\sum_{1\le i_1<\cdots<i_k<j} u\left(x\left[\bigcap_{s=1}^{k} I_{i_s} \cap I_j\right]\right) \\
&= \sum_{j=1}^{m} u(x[I_j]) + \sum_{k=1}^{m-1}(-1)^k \sum_{1\le i_1<\cdots<i_k<i_{k+1}\le m} u\left(x\left[\bigcap_{s=1}^{k+1} I_{i_s}\right]\right) \\
&= \sum_{j=1}^{m} u(x[I_j]) + \sum_{k=2}^{m}(-1)^{k+1} \sum_{1\le i_1<\cdots<i_k\le m} u\left(x\left[\bigcap_{s=1}^{k} I_{i_s}\right]\right) \\
&= \sum_{k=1}^{m}(-1)^{k+1} \sum_{1\le i_1<\cdots<i_k\le m} u\left(x\left[\bigcap_{s=1}^{k} I_{i_s}\right]\right) \\
&= u(x) \qquad \text{by (11.8)},
\end{aligned}$$

from which (11.7) readily follows. ◆

11.4 PROBABILITY MEASURES ON HOMOGENEOUS PRODUCT SETS

Throughout this section $X = A^n$, $\mathcal{P}_s$ is the set of simple probability measures on X, and $\mathcal{R}$ is the set of simple probability measures on A. For $P \in \mathcal{P}_s$, $P_i \in \mathcal{R}$ is the marginal measure of P on the ith A: that is, $P_i(B) = P(A^{i-1} \times B \times A^{n-i})$ for $B \subseteq A$. The marginal measure of P on all but the ith A will be denoted P_i^c: $P_i^c((x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)) = \sum_{a\in A} P((x_1, \ldots, x_{i-1}, a, x_{i+1}, \ldots, x_n))$.

Based on $<$ on $\mathcal{P}_s$, we define $<$ on $\mathcal{R}$ as follows:

$R < R^* \Leftrightarrow P < Q$ for every $P, Q \in \mathcal{P}_s$ such that $P_i = R$ and $Q_i = R^*$ for all i. Three special preference conditions will be applied to this case:

C1. $[P, Q \in \mathcal{P}_s, P_i = Q_i$ *for* $i = 1, \ldots, n] \Rightarrow P \sim Q$.

C2. $[P, Q \in \mathcal{P}_s, P_i = R, Q_i = R^*, P_i^c = Q_i^c] \Rightarrow [P < Q \Leftrightarrow R < R^*]$.

C3. *For some* $R \in \mathcal{R}$, $[P, Q, P^*, Q^* \in \mathcal{P}_s, P_n = Q_n = P_1^* = Q_1^* = R, P_n^c = Q_n^c, P_1^{*c} = Q_1^{*c}] \Rightarrow [P < Q \Leftrightarrow P^* < Q^*]$.

C2 is a persistence condition, much like the definition of persistence in Section 7.1. Under C1, all P that have $P_i = R$ for $i = 1, \ldots, n$ are indifferent, and all Q that have $Q_i = R^*$ for all i are indifferent. Hence if $<$ on $\mathcal{P}_s$ is a weak order then, if $P < Q$ for one such P and Q, $P < Q$ for all such P and Q so that $<$ on $\mathcal{R}$ is a "faithful" weak order. C2 says that this weak order on $\mathcal{R}$ applies to each of the n factors. C3 is a form of stationary condition, and compares with stationarity as defined in Definition 7.3 of Section 7.3.

The reasonableness of these conditions is, of course, doubtful in most situations.

THEOREM 11.4. *Suppose that there is a real-valued function u on* $X = A^n$ *that satisfies* $P < Q \Leftrightarrow E(u, P) < E(u, Q)$, *for all* $P, Q \in \mathcal{P}_s$, *that* $P < Q$ *for some* $P, Q \in \mathcal{P}_s$, *and that C1 and C2 hold. Then there is a real-valued function* ρ *on A and positive numbers* $\lambda_1, \ldots, \lambda_n$ *such that*

$$P < Q \Leftrightarrow \sum_{i=1}^{n} \lambda_i E(\rho, P_i) < \sum_{i=1}^{n} \lambda_i E(\rho, Q_i), \qquad \text{for all } P, Q \in \mathcal{P}_s, \quad (11.9)$$

and ρ' *on A and positive* $\lambda_1', \ldots, \lambda_n'$ *satisfy* (11.9) *along with* ρ *and* $\lambda_1, \ldots, \lambda_n$ *if and only if there are real numbers* $p > 0, q > 0$ *and r such that*

$$\lambda_i' = p\lambda_i \qquad \text{for } i = 1, \ldots, n \quad (11.10)$$

$$\rho'(a) = q\rho(a) + r \qquad \text{for all } a \in A. \quad (11.11)$$

If, in addition, C3 holds and $n \geq 2$ *then there is a unique number* $\pi > 0$ *such that*

$$P < Q \Leftrightarrow \sum_{i=1}^{n} \pi^{i-1} E(\rho, P_i) < \sum_{i=1}^{n} \pi^{i-1} E(\rho, Q_i), \qquad \text{for all } P, Q \in \mathcal{P}_s. \quad (11.12)$$

Expression (11.9) compares with (7.9) and (11.12) compares with (7.13).

Proof. To obtain (11.9) we use Theorem 11.1 to obtain (11.1) for all $P, Q \in \mathcal{P}_s$, where each u_i is defined on A. $C1$ is used in this. It then follows from $C2$ and the definition of $<$ on $\mathfrak{R}$ that, for each i, $R < R^* \Leftrightarrow E(u_i, R) < E(u_i, R^*)$ for all $R, R^* \in \mathfrak{R}$. It follows from Theorem 8.4 that the u_i are related by positive linear transformations, say $u_j = a_j u_1 + b_j$ with $a_j > 0$ for $j = 2, \ldots, n$. Let $\rho \equiv u_1$ and $\lambda_1 = 1$, $\lambda_j = a_j$ for $j = 2, \ldots, n$. Then (11.9) follows.

Suppose ρ' and $\lambda_i' > 0$ satisfy (11.9) also. Then, since the $\lambda_i \rho$ are unique up to similar positive linear transformations by Theorem 11.1, there are numbers $k > 0$ and $\beta_1, \ldots, \beta_n$ such that $\lambda_i' \rho' = k\lambda_i \rho + \beta_i$ for $i = 1, \ldots, n$. (11.10) and (11.11) then follow as in the proof of Theorem 7.4. $P < Q$ for some $P, Q \in \mathcal{P}_s$ is used in obtaining (11.10).

The proof for (11.12) follows the general lines given in the proof of Theorem 7.5 and will not be detailed here. ◆

11.5 SUMMARY

When $X \subseteq \prod_{i=1}^{n} X_i$, the usual expected utility axioms along with a condition that says that $P \sim Q$ when the marginal measure of P for X_i equals the marginal measure of Q for X_i $(i = 1, \ldots, n)$ leads to the additive

form $P \prec Q \Leftrightarrow \sum_i E(u_i, P_i) < \sum_i E(u_i, Q_i)$. This was proved in general for $X = \prod_{i=1}^n X_i$ and for $X \subseteq X_1 \times X_2$. (It is true also for simple measures when $X \subseteq \prod_{i=1}^n X_i$, but the proof of this was discovered too late for inclusion here.)

Under the additive, expected utility representation in the homogeneous context with $X = A^n$, a persistence condition leads to $P \prec Q \Leftrightarrow \sum_i \lambda_i E(\rho, P_i) < \sum_i \lambda_i E(\rho, Q_i)$, and persistence and stationarity lead to $P \prec Q \Leftrightarrow \sum_i \pi^{i-1} E(\rho, P_i) < \sum_i \pi^{i-1} E(\rho, Q_i)$.

INDEX TO EXERCISES

1. 50–50 indifference condition. 2. Binary relations on the $\mathcal{P}_i$. 3. Marginal expectations. 4. Additivity with finite $X \subseteq \Pi X_i$. 5–8. Alternating sequences and cycles. 9. Markovian dependence in utility theory. 10. Admissible transformations. 11. Theorem 11.3 versus Theorem 11.2. 12. No time preference. 13–14. Theorem 11.4.

Exercises

1. For the bridge-construction example of Section 10.1 let x_1 be cost and let x_2 be completion time in $(x_1, x_2) \in X_1 \times X_2$. Assume that both factors are subject to uncertainty. With $X = X_1 \times X_2$ argue that only 50-50 gambles of the following form need to be used in testing the indifference condition of Theorem 11.1: P gives (\$100 million, 4 years) or (x_1 million, x_2 years) each with probability .5; Q gives (\$100 million, x_2 years) or (x_1 million, 4 years) each with probability .5.

2. Let $\mathcal{P}$ be the set of simple probability measures on $X = \prod_{i=1}^n X_i$ and let $\mathcal{P}_i$ be the set of simple probability measures on X_i. With $a, b \in \mathcal{P}_i$ define $a \precsim_i b \Leftrightarrow P \precsim Q$ for every $P, Q \in \mathcal{P}$ such that $P_i = a$, $Q_i = b$, and $P_i^c = Q_i^c$, where $P_i^c(Q_i^c)$ is the marginal of $P(Q)$ on $\prod_{j \neq i} X_j$. Also let $a \prec_i b \Leftrightarrow (a \precsim_i b$, not $b \precsim_i a)$, $a \sim_i b \Leftrightarrow (a \precsim_i b, b \precsim_i a)$. We identify the following conditions:

A. $\precsim$ on $\mathcal{P}$ is transitive and connected;
B. $(P \prec Q, 0 < \alpha < 1) \Rightarrow \alpha P + (1 - \alpha)R \prec \alpha Q + (1 - \alpha)R$;
C. $(P \sim Q, 0 < \alpha < 1) \Rightarrow \alpha P + (1 - \alpha)R \sim \alpha Q + (1 - \alpha)R$;
D. $(P_i = Q_i$ for $i = 1, \ldots, n) \Rightarrow P \sim Q$,

where $P \prec Q \Leftrightarrow (P \precsim Q$, not $Q \precsim P)$ and $P \sim Q \Leftrightarrow (P \precsim Q, Q \precsim P)$. Prove the following theorems. The $\not\Rightarrow$ means "does not imply."

a. ($\precsim$ on $\mathcal{P}$ is transitive) $\Rightarrow$ ($\precsim_i$ on $\mathcal{P}_i$ is transitive).
b. $(P_i \sim_i Q_i, P_i^c = Q_i^c) \Rightarrow P \sim Q$.
c. $(P_i \prec_i Q_i, P_i^c = Q_i^c) \not\Rightarrow P \prec Q$.
d. ($\precsim$ is transitive, $P_i \precsim_i Q_i$ for all i) $\Rightarrow P \precsim Q$.
e. (A, each $\precsim_i$ on $\mathcal{P}_i$ is transitive and connected, $P_i \precsim_i Q_i$ for all i, $P_i \prec Q_i$ for some i) $\not\Rightarrow P \prec Q$.
f. $D \Rightarrow \precsim$ on $\mathcal{P}$ and $\precsim_i$ on $\mathcal{P}_i$ are reflexive.

g. ($n = 2$, $\preccurlyeq_i$ is reflexive for $i = 1, 2) \Rightarrow D$.
h. ($n > 2$, $\preccurlyeq_i$ is reflexive for each i) $\not\Rightarrow D$.
i. ($n \geq 2$, $\preccurlyeq_i$ is reflexive for each i, $\preccurlyeq$ is transitive) $\Rightarrow D$.
j. $(A, D) \not\Rightarrow \preccurlyeq_i$ on $\mathcal{P}_i$ is transitive and connected.
k. ($\preccurlyeq$ is transitive, $B, C, D, P_i = a, Q_i = b, P_i^c = Q_i^c, a \prec_i b) \Rightarrow P \prec Q$.
l. ($\preccurlyeq$ is transitive, $B, C, D, P_i \preccurlyeq_i Q_i$ for all i, $P_i \prec_i Q_i$ for some $i) \Rightarrow P \prec Q$.
m. When C in k and l is replaced by A, the conclusions of the two theorems can be false.
n. $(A, B, C, D) \Rightarrow \preccurlyeq_i$ on $\mathcal{P}_i$ is transitive and connected.
o. $(A, B, C, D, P_i \preccurlyeq_i Q_i, 0 < \alpha < 1) \Rightarrow \alpha P_i + (1 - \alpha)R_i \preccurlyeq_i \alpha Q_i + (1 - \alpha)R_i$.
p. $(A, B, C, D, P_i \prec_i Q_i, 0 < \alpha < 1) \Rightarrow \alpha P_i + (1 - \alpha)R_i \prec_i \alpha Q_i + (1 - \alpha)R_i$.

3. With $X \subseteq \prod_{i=1}^n X_i$ let P_i be the marginal measure on X_i of the probability measure P on X and let f on X and f_i on X_i be real-valued functions that satisfy $f(x_1, \ldots, x_n) = f_i(x_i)$ for all $x \in X$. Prove:
a. (P is simple, $X = \prod X_i) \Rightarrow E(f, P) = E(f_i, P_i)$.
b. (P is simple, $X \subseteq \prod X_i) \Rightarrow E(f, P) = E(f_i, P_i)$.
c. (f_i is bounded, $X \subseteq \prod X_i) \Rightarrow E(f, P) = E(f_i, P_i)$.

4. *Suppose that* $\mathcal{P}$ *is the set of simple probability measures on a finite set* $X \subseteq \prod_{i=1}^n X_i$, *that there is a real-valued function* u *on* X *that satisfies* $P \prec Q \Leftrightarrow E(u, P) < E(u, Q)$ *for every* $P, Q \in \mathcal{P}$, *and that* $(P, Q \in \mathcal{P}, P_i = Q_i$ *for* $i = 1, \ldots, n) \Rightarrow P \sim Q$. *Then there are real-valued functions* $u_1, \ldots, u_n$ *on* $X_1, \ldots, X_n$ *respectively that satisfy* (11.1).

Prove this theorem using the following steps.
a. To establish $u(x_1, \ldots, x_n) = \sum u_i(x_i)$ for each $x \in X$ note that this system of equations is the same as

$$\sum_{k=1}^{N} a_{jk} v(y_k) = u(x^j) \qquad j = 1, \ldots, M \tag{11.13}$$

when we let $X = \{x^1, \ldots, x^M\}$, $X_i = \{x_{i1}, \ldots, x_{im_i}\}$, $(y_1, \ldots, y_N) = (x_{11}, \ldots, x_{1m_1}, x_{21}, \ldots, x_{n1}, \ldots, x_{nm_n})$, $N = \sum_{i=1}^n m_i$, and define the $a_{jk} \in \{0, 1\}$ in an appropriate manner with $\sum_{k=1}^N a_{jk} = n$ for each j.
b. It is a well-known fact of linear algebra that (11.13) has a v-solution if and only if for any non-zero vector $(c_1, \ldots, c_M) \in \mathrm{Re}^M$

$$\left(\sum_{j=1}^{M} c_j a_{jk} = 0 \text{ for } k = 1, \ldots, N\right) \Rightarrow \sum_{j=1}^{M} c_j u(x^j) = 0. \tag{11.14}$$

To verify this for a non-zero $(c_1, \ldots, c_M)$ let $A = \{j : c_j > 0\}$, $B = \{j : c_j < 0\}$, $P = \sum_A (c_j / \sum_A c_j) x^j$ and $Q = \sum_B (c_j / \sum_B c_j) x^j$, and show that the left side of (11.14) implies that $A \neq \varnothing$ and $B \neq \varnothing$, that $P_i = Q_i$ for $i = 1, \ldots, n$, and that $\sum_A c_j = -\sum_B c_j$. Then use the indifference condition to establish (11.14).

5. In the $X \subseteq X_1 \times X_2$ context of Section 11.2 verify:
a. R in Definition 11.2 is an equivalence.
b. Lemma 11.1. (Consider a shortest sequence.)
c. A cycle has at least four elements.

6. Let $X = \{(x_1, x_2), (x_1, z_2), (y_1, x_2), (y_1, y_2), (z_1, y_2), (z_1, z_2)\}$, and let u on X

satisfy $P \prec Q \Leftrightarrow E(u, P) < E(u, Q)$ for a probability measures P, Q on X. Prove:

a. X is a cycle.

b. X has no four-element cycle.

c. The 50-50 indifference condition of Theorem 11.1 holds.

d. (11.1) can be false.

7. Prove Lemma 11.2 by showing that if D has more than one alternating sequence from (x_1, x_2) to (y_1, y_2) then D includes a cycle.

8. Prove Lemma 11.3.

9. In the context of Section 11.3 let $I_i = \{i, i + 1\}$ for $i = 1, 2, \ldots, n - 1$ with $m = n - 1$.

a. Verify that the indifference condition at the end of Theorem 11.3 implies the following: $[\{(x_i, x_{i+1}), (y_i, y_{i+1})\} = \{(z_i, z_{i+1}), (w_i, w_{i+1})\}$ for $i = 1, \ldots, n - 1] \Rightarrow \frac{1}{2}x + \frac{1}{2}y \sim \frac{1}{2}z + \frac{1}{2}w$. (The latter are 50-50 gambles.)

b. Prove that Theorem 11.3 is true for the case at hand when the indifference condition in Theorem 11.3 is replaced by the 50-50 indifference condition in (*a*). Obtain $u(x) = \sum_{i=1}^{n-1} u_i(x_i, x_{i+1})$.

c. With $P \prec Q \Leftrightarrow E(u, P) < E(u, Q)$ for all $P, Q \in \mathcal{P}$, suppose that $u(x) = \sum_{i=1}^{n-1} u_i(x_i, x_{i+1}) = \sum_{i=1}^{n-1} v_i(x_i, x_{i+1})$ for all $x \in X$. Show that there are real-valued functions $f_2, \ldots, f_{n-1}$ on $X_2, \ldots, X_{n-1}$ such that

$$v_1(a, b) = u_1(a, b) + f_2(b) \qquad \text{for all } (a, b) \in X_1 \times X_2,$$

$$v_i(a, b) = u_i(a, b) - f_i(a) + f_{i+1}(b) \qquad \text{for all } (a, b) \in X_i \times X_{i+1},$$

$$2 \le i \le n - 2,$$

$$v_{n-1}(a, b) = u_{n-1}(a, b) - f_{n-1}(a) \qquad \text{for all } (a, b) \in X_{n-1} \times X_n.$$

10. In the context of Section 11.3 suppose that $u(x) = \sum_{j=1}^{m} u_j(x^j)$ for all $x \in \prod_{i=1}^{n} X_i$, as in the proof of Theorem 11.3. With u fixed, describe the set of transformations on the u_j that preserve equality. Note that, if $\{v_j\}$ is such a transformation of $\{u_j\}$ then $\sum u_j(x^j) = \sum v_j(x^j)$ and consequently $\sum_{k=1}^{m} v_k([I_j]^k) = \sum_{k=1}^{m} u_k([I_j]^k)$ for $j = 1, \ldots, m$ so that

$$v_j(x^j) = u_j(x^j) + \sum_{k \neq j} [u_k([I_j]^k) - v_k([I_j]^k)].$$

If $I_j \cap I_k = \varnothing$, argue that $u_k([I_j]^k) - v_k([I_j]^k)$ is constant as x ranges over X, and if $I_j \cap I_k \neq \varnothing$ then the stated difference varies as x ranges over X but the variation is caused only by the x_i for $i \in I_j \cap I_k$.

11. Argue that if the generalization of Theorem 11.2 were true for $X \subseteq \prod_{i=1}^{n} X_i$ with $n > 2$, then Theorem 11.3 for simple measures would be an immediate corollary of the more general form of Theorem 11.2.

12. Show that the hypotheses in the first two lines of Theorem 11.4 along with $(P_1, \ldots, P_n$ is a permutation of $Q_1, \ldots, Q_n) \Rightarrow P \sim Q$, imply that there is a real-valued function ρ on X such that, for all $P, Q \in \mathcal{P}_s$, $P \prec Q \Leftrightarrow \sum_i E(\rho, P_i) < \sum_i E(\rho, Q_i)$.

13. Verify (11.12) in Theorem 11.4.

14. Can you imagine a situation in the context of Section 11.4 where any one of $C1$, $C2$, and $C3$ seems reasonable with $n > 1$?

PART

III

STATES OF THE WORLD

Preference structures that incorporate uncertainty in the formulation of alternatives but do not presuppose probability have been expressed mostly in states of the world models. In such a model the uncertainty concerns which state in a set of mutually exclusive states (or environments) obtains, or is the "true state." It is generally assumed that (1) the decision maker does not know the "true state," (2) the act he selects has no effect on the state that obtains, and (3) the state that obtains affects the outcome of the decision in conjunction with the act selected.

Interest in expected-utility theories that are set in the states of the world formulation is due in large part to Leonard J. Savage's theory (Chapter 14), published in 1954. Before this, the now widely-referenced theory of Frank P. Ramsey (1931) was virtually unknown. Savage's theory reflects elements from Ramsey and from John von Neumann and Oskar Morgenstern: his interpretation of probability owes much to the pioneering work of Bruno de Finetti.

Some other theory for the states formulation is presented in Chapters 12 and 13.

Chapter 12

STATES OF THE WORLD

This chapter introduces the states of the world formulation for decision under uncertainty. The first section describes the usual states formulation and compares it with the approach of Part II. The second section examines the weak-order expected-utility model for the states formulation, and discusses several axiomatic approaches to the model. Several of these approaches are explored in the next two chapters.

The second section also points out two problems that arise in the theories. One of these, often referred to as the "constant acts" problem, suggests an alternative approach to the expected-utility model. Axioms for the alternative approach have yet to be discovered. The second problem concerns the fineness of state descriptions and residual uncertainty. Some additive utility models that are designed for this possibility and which do not explicitly include state probabilities are discussed in the third section.

12.1 STATES AND STATES

In Part II of this book we thought of a decision under uncertainty in terms of a set F of available acts or strategies and a set X of consequences, one of which will follow from the selected act. We assumed that the decision maker's uncertainty about which $x \in X$ would occur if $f \in F$ were selected could be expressed by a probability measure P_f on X. The axioms were based on sets of probability measures that supposedly included $\{P_f : f \in F\}$.

To enlarge on this let S' be the set of functions on acts to consequences. Each $s \in S'$ assigns a consequence $s(f) \in X$ to each $f \in F$. Suppose, for example, that a young man will propose marriage to either Alice or Betsy, but not both in case one refuses him. Suppose further that he is interested only in the three consequences in {Marry Alice, Marry Betsy, Stay Single}. In this case S' contains nine functions, but only four of these need be considered. The four are {(Propose to Alice, Marry Alice), (Propose to Betsy, Marry Betsy)}, {(Propose to Alice, Marry Alice), (Propose to Betsy, Stay

Single)}, {(Propose to Alice, Stay Single), (Propose to Betsy, Marry Betsy)}, and {(Propose to Alice, Stay Single), (Propose to Betsy, Stay Single)}. One of the five functions that is excluded is {(Propose to Alice, Marry Betsy), (Propose to Betsy, Marry Alice)}.

Suppose that if act f is implemented then consequence $s(f)$ will occur and this is true for each $f \in F$. Then we say that s *obtains*. By $C' \subseteq S'$ obtains we mean that some $s \in C'$ obtains. Suppose the decision maker has a probability measure P' on (the set of subsets of) S'. $P'(C')$ is interpreted as a measure of his belief in the truth of the proposition "C' obtains." Given P' we could define P_f by

$$P_f(A) = P'(\{s : s \in S', s(f) \in A\}) \qquad \text{for all } A \subseteq X. \tag{12.1}$$

In the marriage example we would expect that $P'(C') = 0$ when C' is the set of the five "excluded" functions. Then we would have P' (either girl would say "yes") + P' (only Alice would say "yes") + P' (only Betsy would say "yes") + P' (neither girl would say "yes") = 1. Here we have translated the four functions into conditions under which they will obtain. For example {(Propose to Alice, Marry Alice), (Propose to Betsy, Stay Single)} obtains if and only if only Alice would say "yes."

If $s \in S'$ obtains it will obtain regardless of which f is implemented. This is a result of the way S' has been formulated. Hence the decision maker's choice should not influence his beliefs about which s might obtain. But we expect that his beliefs about which s might obtain will influence his choice.

In most cases P' on S' contains more information about the decision maker's uncertainty than does $\{P_f : f \in F\}$ when the P_f are probability measures defined from P' as in (12.1). To determine an act in F that maximizes expected utility it is usually unnecessary to estimate all of P', a task that may be an order of magnitude more difficult than the estimation of the P_f.

Although four potentially nonzero $P'(s)$ were noted in the marriage example, our young man would presumably be satisfied with estimating the two probabilities $p = P_{\text{propose to Alice}}$ (Marry Alice) $= P_{\text{propose to Alice}}$ (Alice would say "yes") and $q = P_{\text{propose to Betsy}}$ (Marry Betsy) $=$ $P_{\text{propose to Betsy}}$ (Betsy would say "yes"). In fact, all he needs is an estimate of the ratio p/q since $E(u,$ Propose to Alice$) < E(u,$ Propose to Betsy$) \Leftrightarrow$ $p/q < [u(\text{Marry Betsy}) - u(\text{Stay Single})]/[u(\text{Marry Alice}) - u(\text{Stay Single})]$.

States of the World

In Savage's words (1954, p. 9) the *world* is "the object about which the person is concerned" and a *state* of the world is "a description of the world, leaving no relevant aspect undescribed." The states are to incorporate all decision-relevant factors about which the decision maker is uncertain and should be formulated in such a way that the state that obtains does not depend on the act selected.

According to the last part of this description it would not seem out of place to call the elements in S' "states." However, the approach made popular by Savage and others does not usually proceed in this way. Instead of defining states as functions on acts to consequences, Savage defines acts as functions on states to consequences. With S the set of states of the world, each $f \in F$ is a function on S to X: $f(s)$ is the consequence that occurs if f is implemented and $s \in S$ obtains.

Simple examples of states as they are often thought of in the Savage approach are: whether an unbroken egg (the world) is good (state 1) or rotten (state 2); whether the next flip of this coin will result in a head (s_1) or a tail (s_2); whether the accused is guilty (s_1) or innocent (s_2); whether these mushrooms are harmless (s_1) or poisonous (s_2). If $f =$ "Eat the mushrooms" and $g =$ "Throw away the mushrooms" then $f(s_1) =$ "Enjoy a culinary treat," $f(s_2) =$ "Enjoy a culinary treat then die," and $g(s_1) = g(s_2) =$ "Throw away the bunch of mushrooms."

If S is so formulated that at most one $s \in S$ can obtain, the decision maker cannot conceive of none of them obtaining, and the state that obtains does not depend on the act selected, then we might suppose that the decision maker has a probability measure P^* on S where $P^*(C)$ is his probability that some $s \in C$ with $C \subseteq S$ obtains. We would then define P_f by

$$P_f(A) = P^*(\{s : s \in S, f(s) \in A\}) \qquad \text{for all } A \subseteq X. \tag{12.2}$$

If in fact subsets of S are more or less probable depending on which $f \in F$ is chosen, then new states defined as functions on F into S will remove this difficulty. In most discussions based on Savage's theory it is presumed that (12.2) holds.

Comparisons of Two Formulations

The rest of this book is primarily concerned with utility theory based on Savage's conception of decision under uncertainty. Before we get into that it seems advisable to note that the two formulations presented above are not incompatible. In fact, they are virtually isomorphic when a certain consistent way of viewing uncertainties is adopted. A demonstration of this follows.

Whether S' and S as conceived of above appear different, at least on the surface, suppose in fact that their probability measures P' and P^* agree with each other. By this we mean that, for any $A \subseteq X$ and $f \in F$,

$$P'(\{s' : s' \in S', s'(f) \in A\}) = P^*(\{s : s \in S, f(s) \in A\}). \tag{12.3}$$

This says that the decision maker's probability of getting an $x \in A$ when f is used is independent of the particular method used to describe his uncertainty.

Let u on X be the point utility function defined in such a way that for any two measures P and Q on X, $P < Q \Leftrightarrow E(u, P) < E(u, Q)$. We assume (see Chapter 10 if X is infinite) that u on X is bounded. Let $u_1, u_2, \ldots$ be a sequence

of simple functions on X that converges uniformly from below to u (Definition 10.11). Consider one of these, say u_n. Let u_n have m values with $u_n(A_i) = c_i$ for $i = 1, \ldots, m$ where $\{A_1, \ldots, A_m\}$ is a partition of X and let

$$C_i' = \{s' : s' \in S', s'(f) \in A_i\}, \qquad C_i = \{s : s \in S, f(s) \in A_i\}.$$

Then $\{C_1', \ldots, C_m'\}$ and $\{C_1, \ldots, C_m\}$ are partitions of S' and S respectively and, by (12.3), $\sum_i c_i P'(C_i') = \sum c_i P^*(C_i)$. It follows from Definition 10.12 that

$$E[u(\hat{s}'(f)), P'] = E[u(f(\hat{s})), P^*], \tag{12.4}$$

where $\hat{}$ denotes the varying factor under P' or P^*. In terms of (12.1) the left side of (12.4) is $E(u, P_f)$. In terms of (12.2) the right side of (12.4) is $E(u, P_f)$.

Hence, under the agreement of (12.3), the two formulations give the same value for the expected utility of act f.

12.2 EXPECTED UTILITY PREVIEW

In viewing acts as functions on states to consequences, we shall be concerned with the expected-utility model

$$f < g \Leftrightarrow E[u(f(s)), P^*] < E[u(g(s)), P^*], \qquad \text{for all } f, g \in F, \tag{12.5}$$

where P^* is a probability measure on the set of all subsets of S and u is a utility function on X.

A number of axiomatizations of (12.5) have been made. By far the best known of these is Savage's theory (1954), all of whose axioms can be stated in terms of $<$ on F. His axioms require, among other things, that S be infinite and that if $B \subseteq S$ and $0 \leq \rho \leq 1$ then $\rho P^*(B) = P^*(C)$ for some $C \subseteq B$. He assumes also that every element in X can occur under each state and that all constant acts—those that assign the same consequence to every state—are in F. His reason for doing this is to provide a way of defining preferences among consequences on the basis of preferences among (constant) acts. Moreover, this enables the derivation of a probability measure P^* on S. Savage's theory will be presented in detail in Chapter 14.

One of the most criticized aspects of Savage's theory is the structural condition that all of X is relevant under each $s \in S$. For the general situation let $X(s)$ denote the subset of consequences that might actually occur under the acts in F if s obtains. Then, viewing the consequences as complete descriptions of what might occur, it would not seem unusual to have $X(s) \cap X(s') = \varnothing$ when $s \neq s'$. When this is so, there is no natural way of defining preferences on consequences in terms of preferences on acts. This suggests an alternative approach to (12.5) that is based on a pair of preference relations, $<$ on F and $<'$ on X. In this approach we would be interested in

conditions for $<$ and $<'$ that imply the existence of a real-valued function u on $X = \bigcup_S X(s)$ and a probability measure P^* on S that satisfy (12.5) along with $x <' y \Leftrightarrow u(x) < u(y)$, for all $x, y \in X$. I do not presently know of any axiomatization that does this, even when X and F are finite, and allows for no overlap of the $X(s)$.

Extraneous Probabilities

In addition to Savage's approach to (12.5), a number of authors have developed theories that use a set of extraneous probabilities in the axioms. These probabilities may have nothing to do with P^* which, like u, is to be derived from the axioms. Conceptually, the extraneous probabilities can be associated with the outcomes of chance devices such as roulette wheels, dice, or pointers spun on circular disks. The axioms in these cases apply $<$ to a set of elements constructed from F and the extraneous probabilities. The set to which $<$ is applied includes F as a special subset.

Axioms for (12.5) that use extraneous probabilities from 0 to 1 have been presented by Chernoff (1954), Anscombe and Aumann (1963), Pratt, Raiffa, and Schlaifer (1964), Arrow (1966), and Fishburn (1969). The next chapter examines two versions of this theory. The first, which assumes that S is finite, follows Pratt, Raiffa, and Schlaifer and assumes only a minimal overlap among the $X(s)$ for different $s \in S$. This overlap is necessary in order to have a base on which to define P^*. The second theory makes no restrictions on the sizes of S and X, but it does assume that $X(s) = X$ for all s as in Savage's theory. However, unlike Savage's theory, almost no restrictions are placed on P^*.

Axioms for (12.5) that use only the extraneous probability 1/2 (or the notion of even-chance gambles) have been developed by Suppes (1956). Suppes' theory can be viewed as a logical completion of Ramsey's (1931) ideas. Suppes (1956) should be consulted for a more detailed account. Some of the 50-50 theory is presented in the exercises of Chapter 13.

Residual Uncertainty and Act-State Pairs

In practice it is seldom possible to ensure that the states will leave no relevant aspect of the world undescribed. No matter how finely we describe the potential realizations of the world, the descriptions will usually be incomplete even when the states meet the logical criteria of being mutually exclusive and collectively exhaustive. Thus the specification of act f and state s will enable us to say something about what will occur although we may never be precisely certain about exactly what will happen if f is implemented and s obtains. Part of this residual uncertainty can be identified explicitly by expanding S to obtain a finer set of states. This may necessitate an expansion of F also.

The practical question is thus seen as the question of how detailed to make the states in light of the purpose of the decision and the import of the potential consequences.

The possibility of residual uncertainty (given f and s we are still not precisely certain about what will happen) leads us to consider a formulation that does not attempt to detail exact consequences. In this formulation consequences $f(s)$ are replaced by act-state pairs $(f, s) \in F \times S$. Uncertainties not resolved by simply specifying act-state pairs might be mentally factored into the situation by the decision maker during his preference deliberations.

In this case no act-state pair appears under more than one state. Thus we have the kind of situation described above where $X(s) \cap X(s') = \varnothing$. With u a utility function on $F \times S$ in the present formulation, we might ask for conditions for a binary relation $<$ on F and a binary relation $<'$ on $F \times S$ that imply the existence of a real-valued function u on $F \times S$ and a probability measure P^* on S such that

$$(f, s) <' (g, t) \Leftrightarrow u(f, s) < u(g, t), \qquad \text{for all } (f, s), (g, t) \in F \times S, \quad (12.6)$$

$$f < g \Leftrightarrow E[u(f, s), P^*] < E[u(g, s), P^*], \qquad \text{for all } f, g \in F. \quad (12.7)$$

I do not presently know of any more-or-less satisfactory axiomatization for this model.

12.3 MODELS WITHOUT STATE PROBABILITIES

Despite the absence of axioms for the $F \times S$ model of (12.6) and (12.7) we can formulate axioms for more general but perhaps somewhat less interesting forms of that model. These forms posit additivity over the states but make no attempt to define state probabilities. I shall comment briefly on several of them. These comments apply also to the consequence formulation when $X(s) \cap X(s') = \varnothing$ whenever $s \neq s'$. Throughout this section both F and S are assumed to be finite.

An Order for Each State

In our first case we assume that the decision maker has a weak order $<$ on F along with a weak order $<_s$ on F for each s. Thus, $<_s$ orders F under the hypothesis that s obtains. If the decision maker does in fact have a weak order $<'$ on $F \times S$, then $<_s$ would be obtained as the restriction of $<'$ to $F \times \{s\}$. In this context we are interested in the existence of a real-valued function v on $F \times S$ that satisfies

$$f <_s g \Leftrightarrow v(f, s) < v(g, s), \qquad \text{for all } f, g \in F \text{ and } s \in S, \quad (12.8)$$

$$f < g \Leftrightarrow \textstyle\sum_S v(f, s) < \sum_S v(g, s), \qquad \text{for all } f, g \in F. \quad (12.9)$$

Under the weak order conditions, an independence axiom across states that is necessary and sufficient for (12.8) and (12.9) can be derived from the Theorem of The Alternative (Theorem 4.2). One version of such an axiom is: if $f^j \preccurlyeq g^j$ (i.e., $f^j \prec g^j$ or $f^j \sim g^j$) for $j = 1, \ldots, m$ and if for each s there is a permutation $f^{s1}, \ldots, f^{sm}$ of $f^1, \ldots, f^m$ such that $g^j \preccurlyeq_s f^{sj}$ for $j = 1, \ldots, m$, then in fact $f^j \sim g^j$ and $g^j \sim_s f^{sj}$ for all j and s.

Apart from the question of intransitive indifference, one can criticize the model of (12.8) and (12.9) on the count that the decision maker might have a nonindifferent weak order $\prec_s$ when he regards s as virtually impossible. In (12.7) we can take care of this by setting $P^*(s) = 0$, but the only way of reflecting "s is impossible" in (12.9) in a general way is to have $v(f, s)$ constant on F for the given s, and if (12.8) is to hold we then require $(f, s) \sim_s (g, s)$ for all $f, g \in F$. The model given by (12.8) and (12.9) can easily be amended to handle this criticism by excluding all states that, in the judgment of the decision maker, cannot possibly obtain. Such states will be referred to as *null states* in the next two chapters.

The model and therefore the independence condition can easily be seen to be unreasonable when the "state" that obtains depends on the selected act. For example, suppose you want to sell something and can either advertise (at some cost) or not advertise. Let the "states" be s = item is sold, t = item is not sold. Then surely advertise $\prec_s$ don't advertise, and advertise $\prec_t$ don't advertise. According to the model this requires that advertise $\prec$ don't advertise, which, if we took it seriously, would say that one should never advertise.

Perfect Information Acts

An alternative to using the $\prec_s$ directly is to work only with $\prec$ on a set that includes F. For example, let $\mathcal{F}$ be the set of functions on S to F: a function $\mathfrak{f} \in \mathcal{F}$, which assigns act $\mathfrak{f}(s)$ to state s for each $s \in S$, is a *perfect information act*. We interpret $\mathfrak{f}$ as follows. Suppose the decision maker specifies an $\mathfrak{f} \in \mathcal{F}$. He then gives this function to an imaginary second party who has perfect information about which state obtains and who proceeds to implement the $\mathfrak{f}(s) \in F$ for the s that obtains. In terms of the $\prec_s$ the decision maker's most preferred $\mathfrak{f}$ would presumably be one that for each s has $f \preccurlyeq_s \mathfrak{f}(s)$ for all $f \in F$. (This assumes that the state that obtains does not depend on the selected act.) F is the subset of constant functions in $\mathcal{F}$.

Under this formulation we are interested in the existence of a real-valued function v on $F \times S$ that satisfies

$$\mathfrak{f} \prec \mathfrak{g} \Leftrightarrow \sum_S v(\mathfrak{f}(s), s) < \sum_S v(\mathfrak{g}(s), s), \qquad \text{for all } \mathfrak{f}, \mathfrak{g} \in \mathcal{F}. \tag{12.10}$$

Condition C of Theorem 4.1 applies directly to this case. That is, (12.10)

holds if and only if $[\mathfrak{f}^1(s), \ldots, \mathfrak{f}^m(s)$ is a permutation of $\mathfrak{g}^1(s), \ldots, \mathfrak{g}^m(s)$ for each $s \in S$, $\mathfrak{f}^j \preccurlyeq \mathfrak{g}^j$ for $j = 1, \ldots, m-1] \Rightarrow$ not $\mathfrak{f}^m \prec \mathfrak{g}^m$.

A probabilistic argument (using extraneous probabilities) that supports this independence condition proceeds as follows. Suppose $\mathfrak{f}^1(s), \ldots, \mathfrak{f}^m(s)$ is a permutation of $\mathfrak{g}^1(s), \ldots, \mathfrak{g}^m(s)$ for each s. Let $\sum (1/m)\mathfrak{f}^j$ denote an "alternative" whose "implementation" is carried out as follows. A well-balanced die with m symmetric faces numbered 1 through m is rolled and if face j occurs then $\mathfrak{f}^j$ is used, with $(\mathfrak{f}^j(s), s)$ the resulting act-state pair if s obtains. $\sum (1/m)\mathfrak{g}^j$ has a similar interpretation. Supposing for convenience that all $\mathfrak{f}^j(s)$ are different for $j = 1, \ldots, m$, if s obtains then $\sum (1/m)\mathfrak{f}^j$ gives each of $(\mathfrak{f}^1(s), s), \ldots, (\mathfrak{f}^m(s), s)$ an equal chance of resulting. The same is true with respect to $\sum (1/m)\mathfrak{g}^j$, and since $\mathfrak{g}^1(s), \ldots, \mathfrak{g}^m(s)$ is a permutation of $\mathfrak{f}^1(s), \ldots, \mathfrak{f}^m(s)$ it seems natural to regard $\sum (1/m)\mathfrak{f}^j$ and $\sum (1/m)\mathfrak{g}^j$ as essentially equivalent if s obtains. Since this is true for each s we would expect that $\sum (1/m)\mathfrak{f}^j \sim \sum (1/m)\mathfrak{g}^j$.

Now if in fact the condition is violated by $\mathfrak{f}^j \preccurlyeq \mathfrak{g}^j$ for all $j < m$ and $\mathfrak{f}^m \prec \mathfrak{g}^m$ we would then expect that $\sum (1/m)\mathfrak{f}^j \prec \sum (1/m)\mathfrak{g}^j$, which violates our "reasonable" conclusion that $\sum (1/m)\mathfrak{f}^j \sim \sum (1/m)\mathfrak{g}^j$.

Extraneous Probabilities

The model given by (12.10) can be embedded in a model that uses extraneous probabilities. In particular, let $\mathcal{P}$ be the set of (simple) probability measures on $\mathcal{F}$. A pseudo-operational interpretation for $P \in \mathcal{P}$ is that, using P, an $\mathfrak{f} \in \mathcal{F}$ is determined: then, if s obtains, $(\mathfrak{f}(s), s)$ is the resulting act-state pair. In this formulation the axioms of Theorem 11.1, letting $X_i = F \times \{s_i\}$, lead to

$$P \prec Q \Leftrightarrow \sum_S E[v(f, s), P_s] < \sum_S E[v(f, s), Q_s], \qquad (12.11)$$

in which P_s is the marginal of P on $F \times \{s\}$ and the $v(\cdot, s)$ for the $s \in S$ are unique up to similar positive linear transformations. (12.10) follows from (12.11) when we define $\mathfrak{f} \prec \mathfrak{g} \Leftrightarrow P \prec Q$ when $P(\mathfrak{f}) = Q(\mathfrak{g}) = 1$.

12.4 SUMMARY

The usual states of the world formulation views the acts in F as functions on states to consequences. The states represent a partition of the potential realizations of the world, ideally leaving no relevant aspect undescribed. It is usually assumed that the state of the world that obtains does not depend on the act selected by the decision maker. If this is not true, new states that satisfy the independence criterion can be defined as functions on acts to the initial set of states. This reformulation is similar to the definition of states

as functions on acts to consequences, as suggested in this chapter in connection with the acts-consequences model of Part II. Under a fundamental agreement between the usual states model and the Part II model, the two models are seen to be alternative but equivalent ways to characterize decision under uncertainty with an expected-utility model.

In cases where acts and states are formulated but exact consequences may not be detailed, independence axioms over the states lead to additive utility models that do not explicitly include probabilities for the states.

INDEX TO EXERCISES

1. Conditional consequence probabilities. 2. Equivalence of two approaches. 3. Job-changing example. 4. Psychology of timing. 5. Independence axiom. 6. Win-lose example and state probabilities. 7–8. Penalty kick example. 9. Propose to the other girl. 10. Theorem of The Alternative for (12.8)–(12.9).

Exercises

1. Use (12.1) and (10.5) to write the probability of "f will result in an $x \in A$, given that g will result in an $x \in A'$" in terms of P'. Then use (12.2) to write the probability in terms of P^*.

2. With all sets finite the utility of act f in the Part II approach can be written as $\sum_X u(x)P_f(x)$, and as $\sum_S u(f(s))P^*(s)$ in the states of the world model. Assuming that $P_f(x) = P^*(\{s: f(s) = x\})$, show that $\sum_S u(f(s))P^*(s) = \sum_X u(x)P_f(x)$.

3. A man currently making \$10000 per year has been offered \$14000 per year by another company. He decides to give his company notice that he will quit unless he gets a new salary of \$$x$. He decides to make x either 13000, 14000, or 15000. The higher x is, the more likely his company will be to reject his ultimatum: if they reject, he will take the new job at \$14000. Formulate his decision under the Part II approach. Then reformulate it in the states of the world manner so that the state that obtains doesn't depend on the selected x.

4. Suppose that you have to make a choice between f and g when the pay-off from your choice will depend on the outcome of one flip of a slightly bent coin that you have been shown. Furthermore, you can either (1) select f or g, after which the coin is flipped by a referee or (2) have the referee make the flip before you choose f or g, but be informed of the outcome of the flip only after you have made your choice. Assuming that you believe the referee is thoroughly honest, do you feel that the procedure (1) or (2) that you select will in any way affect your decision between f and g? Explain the reason(s) for your answer.

5. Adapted from Ellsberg (1961) and Raiffa (1961). An urn contains one white ball (W) and two other balls. You know only that the two other balls are either both red (R), or both green (G), or one is red and one is green. Consider the two situations shown below where W, R, and G represent the three "states" (which don't depend on the act selected) according to whether one ball drawn at random is white, red, or green. The dollar figures are what you will be paid after you make your choice and a ball is drawn.

	W	R	G
f	\$100	\$0	\$0
g	\$0	\$100	\$0

	W	R	G
f'	\$100	\$0	\$100
g'	\$0	\$100	\$100

a. Which of f and g do you prefer? Which of f' and g' do you prefer?

b. Show that the pair ($g \prec f, f' \prec g'$) violates the following independence axiom: if $\{f_1(s), f_2(s)\} = \{g_1(s), g_2(s)\}$ for each $s \in S$ and if $f_1 \prec g_1$ then not $f_2 \prec g_2$. Use an argument like that following (12.10) to argue the "inconsistency" of ($g \prec f, f' \prec g'$).

c. If your answers in (*a*) were $g \prec f$ and $f' \prec g'$, does (*b*) convince you that there is something "wrong" with your preferences? Discuss this.

6. Suppose a decision maker can choose one of two strategies, f and g, and his "opponent" can independently choose one of two strategies, f' and g'. Our decision maker is concerned only with the two consequences "win" and "lose." He believes that either might occur for each of the four strategy pairs in $\{f, g\} \times \{f', g'\}$. Eight states, displayed along the top of Figure 12.1, can be used to partition his "world." Each state specifies the strategy chosen by his opponent and, for each of f and g, specifies whether he will win or lose.

a. Does the state that obtains depend on the one of f and g that is chosen by our decision maker?

b. Suppose an additive expected-utility model without state probabilities, similar to that described by (12.11), is used as a basis for estimating the v numbers in the matrix of Figure 12.1. According to this, g is the better act since $3 + 1 < 2 + 3$. In the usual states model, characterized by (12.5), we would have $\sum_1^8 u(f(s_i))P^*(s_i)$ as the expected utility for f and $\sum_1^8 u(g(s_i))P^*(s_i)$ as the expected utility for g with $f(s_i), g(s_i) \in \{\text{win, lose}\}$ for each i. Assuming that these two models agree with one another we should have $P^*(s_2) = 3a$,

	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8
	f'				g'			
	f: win, g: win	f: win, g: lose	f: lose, g: win	f: lose, g: lose	f: win, g: win	f: win, g: lose	f: lose, g: win	f: lose, g: lose
f	0	3	0	0	0	1	0	0
	0	0	2	0	0	0	3	0

Figure 12.1 A v matrix.

$P^*(s_3) = 2a$, $P^*(s_6) = a$, and $P^*(s_7) = 3a$, where $a > 0$. Explain why this is so. What can you tell about $P^*(\{s_1, s_4, s_5, s_8\})$ from the data? Is there any need to estimate $P^*(s_1)$, $P^*(s_4)$, $P^*(s_5)$, and $P^*(s_8)$ when the usual states model is applied?

7. In soccer, a direct penalty kick inside the box can be viewed as a two-person game between the kicker and the opposing goalkeeper. The goalkeeper can select one of three acts:

f = stand firm until the kick is made;
g = move right an instant before the ball is kicked;
h = move left an instant before the ball is kicked.

Assume that the kicker will aim right or left (from goalkeeper's orientation). Assuming several symmetries, the goalkeeper's probabilities are presented in Figure 12.2: β is the probability a goal will not be scored if he moves right and the kick is right. Surely $\beta > \alpha > \gamma$.

	f		g		h	
	kick right	kick left	kick right	kick left	kick right	kick left
Goal prevented	α	α	β	γ	γ	β
Goal scored	$1 - \alpha$	$1 - \alpha$	$1 - \beta$	$1 - \gamma$	$1 - \gamma$	$1 - \beta$

Figure 12.2 Conditional probabilities of consequences.

a. If the goalkeeper considers a right kick and a left kick equally likely (which may of course be false), show that f is best if $2\alpha > \beta + \gamma$ and that either g or h is best if $\beta + \gamma > 2\alpha$.

b. Reformulate this in the typical states model with acts f, g, and h and 16 appropriate states.

8. (*Continuation.*) Suppose in the preceding example we use only the gross states s = kick right and t = kick left and that an estimate of v on $\{f, g, h\} \times \{s, t\}$ in accord with (12.11) gives $v(f, s) = 2$, $v(g, s) = 6$, $v(h, s) = 0$, and $v(f, t) = 3$, $v(g, t) = 0$, $v(h, t) = 9$. According to this, which act is most preferred? Describe the best perfect information act. Do the v's suggest that the goalkeeper considers s more probable than t? Why?

9. Suppose an extraneous probability model like that described by (12.11) gives the following v matrix on $F \times S$ for the marriage example of Section 12.1:

	s_1(both "yes")	s_2(only Alice "yes")	s_3(only Betsy "yes")	s_4(both "no")
Propose to Alice	1	2	0	0
Propose to Betsy	0	0	4	0

a. Which girl would the young man rather marry?

b. Which girl should he propose to? (Which act is preferred?)

c. Suppose that extraneous probabilities are used to scale the young man's utilities on the three consequences, after the theory in Chapter 8, and that u(Stay Single) = 0, u(Marry Betsy) = 3, u(Marry Alice) = 4. (That is, "Marry Betsy" is indifferent to a gamble with probability .75 for "Marry Alice" and probability .25 for "Stay Single.") Argue that this data along with the figures in the above matrix suggest that P^*(Betsy would say "yes") = (14/9) P^* (Alice would say "yes").

10. Use the Theorem of The Alternative to verify that the independence condition following (12.9) along with weak order is sufficient for (12.8)–(12.9).

Chapter 13

AXIOMS WITH EXTRANEOUS PROBABILITIES

This chapter gives two derivations of the expected-utility model

$$f < g \Leftrightarrow E[u(f(s)), P^*] < E[u(g(s)), P^*], \qquad \text{for all } f, g \in F, \quad (13.1)$$

that are based on extraneous probabilities as described in Section 12.2. The first derivation assumes that S is finite and presupposes a minimal overlap of the relevant consequences under each state in S. The second makes no restriction on the size of S but assumes that all consequences are relevant under each state. Both P^* and u are derived from the axioms. Section 13.4 shows how these axioms might apply to the decision model of Part II.

All probability measures in this chapter are defined on the set of all subsets of their basic set. "P^* on S" is an abbreviation for "P^* on the set of all subsets of S."

13.1 HORSE LOTTERIES

The purpose of this section is to define many of the terms used later in this chapter and to prove a theorem for horse lotteries, which are the elements on which $<$ is applied in our axioms.

Throughout the chapter S is the set of states of the world. Subsets of S, called *events*, will be denoted by $\{s\}, A, B, C, \ldots$. A *partition* of S is a set of nonempty events that are mutually exclusive and whose union equals S. A^c is the complement of A in S: $A^c = \{s : s \notin A, s \in S\}$. $\{A, A^c\}$ is a two-part partition of S provided that $\varnothing \subset A \subset S$.

F is the set of acts. Each $f \in F$ is a function on S into X, the set of consequences. $X(s) = \{f(s) : f \in F\}$, the set of consequences under state s. $X = \bigcup_S X(s)$. $\mathcal{P}(s)$ *is the set of simple probability measures* (extraneous) *on* $X(s)$. $\mathcal{P}$ *is the set of simple probability measures on* X.

The phrase "horse lottery" was introduced by Anscombe and Aumann (1963). A *horse lottery* is a function on S that assigns a $P \in \mathcal{P}(s)$ to each $s \in S$.

$\mathcal{H}$ is the *set of all horse lotteries*. Horse lotteries are denoted in bold face as $\mathbf{P}, \mathbf{Q}, \mathbf{R}, \ldots$. We adopt the following pseudo-operational interpretation for $\mathbf{P} \in \mathcal{H}$. If $\mathbf{P}$ is "selected" and $s \in S$ obtains then $\mathbf{P}(s) \in \mathcal{P}(s)$ is used to determine a resultant consequence in $X(s)$.

If $\mathbf{P}, \mathbf{Q} \in \mathcal{H}$ and $0 \le \alpha \le 1$ then $\alpha\mathbf{P} + (1-\alpha)\mathbf{Q}$ is the horse lottery in $\mathcal{H}$ that assigns $\alpha\mathbf{P}(s) + (1-\alpha)\mathbf{Q}(s) \in \mathcal{P}(s)$ to $s \in S$, for each $s \in S$. Under this interpretation, $\mathcal{H}$ is a *mixture set* (Definition 8.3).

Taking $<$ on $\mathcal{H}$ as the basic binary relation, $\mathbf{P} \sim \mathbf{Q} \Leftrightarrow$ (not $\mathbf{P} < \mathbf{Q}$, not $\mathbf{Q} < \mathbf{P}$), and $\mathbf{P} \lesssim \mathbf{Q} \Leftrightarrow (\mathbf{P} < \mathbf{Q}$ or $\mathbf{P} \sim \mathbf{Q})$. Event $A \subseteq S$ is *null* $\Leftrightarrow \mathbf{P} \sim \mathbf{Q}$ whenever $\mathbf{P}(s) = \mathbf{Q}(s)$ for every $s \in A^c$. State s is null $\Leftrightarrow \{s\}$ is null.

The following theorem is similar to Theorem 11.1.

THEOREM 13.1. *Suppose that S is finite and that the following hold for all* $\mathbf{P}, \mathbf{Q}, \mathbf{R} \in \mathcal{H}$:

*A*1. $<$ *on* $\mathcal{H}$ *is a weak order;*

*A*2. $(\mathbf{P} < \mathbf{Q}, 0 < \alpha < 1) \Rightarrow \alpha\mathbf{P} + (1-\alpha)\mathbf{R} < \alpha\mathbf{Q} + (1-\alpha)\mathbf{R}$;

*A*3. $(\mathbf{P} < \mathbf{Q}, \mathbf{Q} < \mathbf{R}) \Rightarrow \alpha\mathbf{P} + (1-\alpha)\mathbf{R} < \mathbf{Q}$ *and* $\mathbf{Q} < \beta\mathbf{P} + (1-\beta)\mathbf{R}$ *for some* $\alpha, \beta \in (0, 1)$. *Then, with* $S = \{s_1, \ldots, s_n\}$, *there are real-valued functions* $u_1, \ldots, u_n$ *on* $X(s_1), \ldots, X(s_n)$ *respectively such that*

$$\mathbf{P} < \mathbf{Q} \Leftrightarrow \sum_{i=1}^{n} E(u_i, \mathbf{P}(s_i)) < \sum_{i=1}^{n} E(u_i, \mathbf{Q}(s_i)), \qquad \textit{for all } \mathbf{P}, \mathbf{Q} \in \mathcal{H}, \quad (13.2)$$

and the u_i that satisfy (13.2) *are unique up to similar positive linear transformations, with u_i constant on $X(s_i)$ if and only if s_i is null.*

Proof. By Theorem 8.4, there is a real-valued function u on $\mathcal{H}$ that satisfies $\mathbf{P} < \mathbf{Q} \Leftrightarrow u(\mathbf{P}) < u(\mathbf{Q})$ and $u(\alpha\mathbf{P} + (1-\alpha)\mathbf{Q}) = \alpha u(\mathbf{P}) + (1-\alpha)u(\mathbf{Q})$ and is unique up to a positive linear transformation when it satisfies these properties. For convenience, we shall write $\mathbf{Q} = (\mathbf{Q}(s_1), \ldots, \mathbf{Q}(s_n)) = (Q_1, \ldots, Q_n)$, with $Q_i \in \mathcal{P}(s_i)$ for each i.

Fix $\mathbf{R} = (R_1, \ldots, R_n)$ in $\mathcal{H}$ and let $\mathbf{P}_i = (R_1, \ldots, R_{i-1}, P_i, R_{i+1}, \ldots, R_n)$. Then, with $\mathbf{P} = (P_1, \ldots, P_n)$, $(1/n)\mathbf{P} + ((n-1)/n)\mathbf{R} = \sum_{i=1}^{n} (1/n)\mathbf{P}_i$. Therefore $u(\mathbf{P}) + (n-1)u(\mathbf{R}) = \sum_{i=1}^{n} u(\mathbf{P}_i)$. Defining u_i on $\mathcal{P}(s_i)$ by

$$u_i(P_i) = u(\mathbf{P}_i) - ((n-1)/n)u(\mathbf{R}),$$

summation over i gives $\sum_{i=1}^{n} u_i(P_i) = \sum_{i=1}^{n} u(\mathbf{P}_i) - (n-1)u(\mathbf{R})$, so that $u(\mathbf{P}) = \sum_{i=1}^{n} u_i(P_i)$.

Let $\mathbf{Q}_i = (R_1, \ldots, R_{i-1}, Q_i, R_{i+1}, \ldots, R_n)$. Then, by the preceding result, $u(\alpha\mathbf{P}_i + (1-\alpha)\mathbf{Q}_i) = u_i(\alpha P_i + (1-\alpha)Q_i) + \sum_{j \ne i} u_j(R_j)$. In addition,

$$\begin{aligned} u(\alpha\mathbf{P}_i + (1-\alpha)\mathbf{Q}_i) &= \alpha u(\mathbf{P}_i) + (1-\alpha)u(\mathbf{Q}_i) \\ &= \alpha u_i(P_i) + (1-\alpha)u_i(Q_i) + \textstyle\sum_{j \ne i} u_j(R_j), \end{aligned}$$

so that $u_i(\alpha P_i + (1-\alpha)Q_i) = \alpha u_i(P_i) + (1-\alpha)u_i(Q_i)$. Since the elements

in $\mathcal{P}(s_i)$ are simple measures, $u_i(P_i) = E(u_i, P_i)$ and (13.2) follows with $u_i(x) = u_i(P_i)$ when $P_i(x) = 1$.

Uniqueness up to similar positive linear transformations follows readily from the uniqueness property for u. If v_i satisfy (13.2) along with the u_i then, with $v(\mathbf{P}) = \sum_{i=1}^{n} E(v_i, P_i)$, $v = au + b$ and $a > 0$. Holding P_j fixed for all $j \neq i$, it then follows that $v_i = au_i + b_i$. This holds for each i.

Clearly, u_i is constant on $X(s_i)$ if and only if s_i is null. ◆

13.2 FINITE STATES THEORY

In order to derive u on $X = \bigcup_S X(s)$ and P^* on S on the basis of Theorem 13.1 when S is finite, two more axioms will be used. The first of these ($A4$) assumes that two consequences x_* and x^* appear in every $X(s)$ and that they are not indifferent. Hence $\{x_*, x^*\} \subseteq X(s) \cap X(t)$ for $s, t \in S$. With a convenient abuse of rigor we shall say that a simple probability measure P that assigns probability 1 to $X(s) \cap X(t)$ is in both $\mathcal{P}(s)$ and $\mathcal{P}(t)$, and write $P \in \mathcal{P}(s) \cap \mathcal{P}(t)$.

The second new axiom ($A5$) is a monotonicity axiom. It says that if s and t are not null then there is the same order under both states for all $P \in \mathcal{P}(s) \cap \mathcal{P}(t)$. In other words, preferences on consequences or simple probability measures on consequences that can occur under different states shall not be state dependent.

THEOREM 13.2. *Suppose that the hypotheses of Theorem* 13.1 *hold and that, in addition,*

A4. There are $x_*, x^* \in X(s)$ for every $s \in S$ such that $\mathbf{P} < \mathbf{Q}$ when $\mathbf{P}(s)[\mathbf{Q}(s)]$ assigns probability 1 *to x_* $[x^*]$ for each $s \in S$;*

A5. If $s, t \in S$, s and t are not null, $P, Q \in \mathcal{P}(s) \cap \mathcal{P}(t)$, and if $\mathbf{P} \in \mathcal{H}$, then ($\mathbf{P}$ with $\mathbf{P}(s)$ replaced by P) $<$ ($\mathbf{P}$ with $\mathbf{P}(s)$ replaced by Q) $\Leftrightarrow$ ($\mathbf{P}$ with $\mathbf{P}(t)$ replaced by P) $<$ ($\mathbf{P}$ with $\mathbf{P}(t)$ replaced by Q).

Then there is a real-valued function u on X and a probability measure P^* on S such that

$$\mathbf{P} < \mathbf{Q} \Leftrightarrow E[E(u, \mathbf{P}(s)), P^*] < E[E(u, \mathbf{Q}(s)), P^*], \qquad \textit{for all } \mathbf{P}, \mathbf{Q} \in \mathcal{H}, \tag{13.3}$$

with $P^*(s) = 0$ if and only if s is null. In addition, if v on X and a probability measure Q^* on S satisfy (13.3) *along with u and P^* then $P^* = Q^*$ and there are numbers $a > 0$ and b such that $v(x) = au(x) + b$ for all $x \in \bigcup_{\{s\,:\,s \text{ not null}\}} X(s)$.*

If we define $<$ on F from $<$ on $\mathcal{H}$ by $f < g \Leftrightarrow \mathbf{P} < \mathbf{Q}$ when $\mathbf{P}(s)[\mathbf{Q}(s)]$ assigns probability 1 to $f(s)[g(s)]$ for each $s \in S$, (13.1) follows immediately from (13.3).

Proof. Let $S = \{s_1, \ldots, s_n\}$. Beginning with the results of Theorem 13.1, $A4$ implies that $I = \{i : s_i \in S$ and s_i is not null$\}$ is not empty. If (13.3) is to hold then $P^*(s_i)u$ must be a positive linear transformation of u_i and hence $P^*(s) = 0 \Leftrightarrow s$ is null.

If $I = \{i\}$, (13.3) follows from (13.2) on setting $P^*(s_i) = 1$ and $u(x) = u_i(x)$ for all $x \in X(s_i)$. u on the rest of X is arbitrary. Clearly, u on $X(s_i)$ is unique up to a positive linear transformation.

If I has more than one element let $\mathcal{P}_{ij}$ denote the set of simple probability measures on $X_{ij} = X(s_i) \cap X(s_j)$ when $i, j \in I$. With a convenient lapse in rigor, take $P \in \mathcal{P}(s_i)$ and $P \in \mathcal{P}(s_j)$ when $P \in \mathcal{P}_{ij}$. By (13.2) and $A5$, $E(u_i, P) < E(u_i, Q) \Leftrightarrow E(u_j, P) < E(u_j, Q)$, for all P, $Q \in \mathcal{P}_{ij}$. Then, by $A4$ and the latter part of Theorem 8.2, there is a unique $r_{ij} > 0$ (invariant under similar positive linear transformations of u_i and u_j) such that

$$u_i(x) - u_i(x_*) = r_{ij}[u_j(x) - u_j(x_*)] \qquad \text{for all } x \in X_{ij}. \tag{13.4}$$

Fix $t \in I$ and define P^* and u by

$$P^*(s_i) = r_{it}\Big/\sum_{i \in I} r_{it} \qquad \text{for all } i \in I$$

$$P^*(s_i) = 0 \qquad \text{for all } i \notin I$$

$$u(x) = [u_i(x) - u_i(x_*)]/P^*(s_i) \qquad \text{when } x \in X(s_i) \quad \text{and} \quad i \in I$$

and $u(x) = 0$ when $x \notin \bigcup_I X(s_i)$. To show that u is well defined we need to prove that

$$u_i(x) - u_i(x_*) = \frac{r_{it}}{r_{jt}}[u_j(x) - u_j(x_*)] \qquad \text{when } i, j \in I \quad \text{and} \quad x \in X_{ij}. \tag{13.5}$$

By (13.4), $r_{it}/r_{jt} = ([u_i(x^*) - u_i(x_*)]/[u_t(x^*) - u_t(x_*)])/([u_j(x^*) - u_j(x_*)]/[u_t(x^*) - u_t(x_*)]) = [u_i(x^*) - u_i(x_*)]/[u_j(x^*) - u_j(x_*)] = r_{ij}$, so that (13.5) follows from (13.4) for i, j. Substitution for u_i into (13.2) then yields (13.3).

It follows easily from $u(x_*) < u(x^*)$ and the uniqueness assertions of Theorem 13.1 that P^* is unique and u on $\bigcup_I X(s_i)$ is unique up to a positive linear transformation. ◆

13.3 HOMOGENEOUS HORSE LOTTERY THEORY

The definitions of Section 13.1 apply to this section.

When S is infinite, the horse-lottery approach meets serious mathematical difficulties if we assume only a minimal overlap of the $X(s)$. Therefore, *we shall assume throughout this section that* $X = X(s)$ *for all* $s \in S$. Some additional definitions follow.

$\mathbf{P}$ is *constant* on event $A \Leftrightarrow \mathbf{P}(s) = \mathbf{P}(t)$ for all s, $t \in A$. When $\mathbf{P}(s) = P$ (in $\mathscr{P}$) for all $s \in A$, we shall say that $\mathbf{P} = P$ on A. $\mathbf{P} = \mathbf{Q}$ on $A \Leftrightarrow \mathbf{P}(s) = \mathbf{Q}(s)$ for each $s \in A$. Thus A is null $\Leftrightarrow \mathbf{P} \sim \mathbf{Q}$ whenever $\mathbf{P} = \mathbf{Q}$ on A^c.

With $P, Q \in \mathscr{P}$, we define $\prec$ on $\mathscr{P}$ on the basis of $\prec$ on $\mathscr{H}$ thus: $P \prec Q \Leftrightarrow \mathbf{P} \prec \mathbf{Q}$ when $\mathbf{P} = P$ and $\mathbf{Q} = Q$ on S. Also, $P \prec \mathbf{Q} \Leftrightarrow \mathbf{P} \prec \mathbf{Q}$ when $\mathbf{P} = P$ on S. $P \sim Q$, $P \sim \mathbf{Q}$, $\mathbf{P} \precsim Q, \ldots$ are defined in similar fashion.

After stating our main theorem we shall prove it by proving a series of lemmas.

THEOREM 13.3. *Suppose that the following axioms hold for all* $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R} \in \mathscr{H}$:

*B*1. $\prec$ *on* $\mathscr{H}$ *is a weak order;*

*B*2. $(\mathbf{P} \prec \mathbf{Q}, 0 < \alpha < 1) \Rightarrow \alpha\mathbf{P} + (1 - \alpha)\mathbf{R} \prec \alpha\mathbf{Q} + (1 - \alpha)\mathbf{R}$;

*B*3. $(\mathbf{P} \prec \mathbf{Q}, \mathbf{Q} \prec \mathbf{R}) \Rightarrow \alpha\mathbf{P} + (1 - \alpha)\mathbf{R} \prec \mathbf{Q} \prec \beta\mathbf{P} + (1 - \beta)\mathbf{R}$ *for some* $\alpha, \beta \in (0, 1)$;

*B*4. $P \prec Q$ *for some* $P, Q \in \mathscr{P}$;

*B*5. (*Event* A *is not null,* $\mathbf{P} = P$ *and* $\mathbf{Q} = Q$ *on* A, $\mathbf{P} = \mathbf{Q}$ *on* A^c) $\Rightarrow$ $(\mathbf{P} \prec \mathbf{Q} \Leftrightarrow P \prec Q)$;

*B*6. $\mathbf{P}(s) \prec \mathbf{R}$ *for all* $s \in S \Rightarrow \mathbf{P} \precsim \mathbf{R}$. $\mathbf{R} \prec \mathbf{Q}(s)$ *for all* $s \in S \Rightarrow \mathbf{P} \precsim \mathbf{Q}$.

Then there is a real-valued function u *on* X *and a probability measure* P^* *on* S *that satisfy* (13.3). *Moreover, when* (13.3) *holds for* u *and* P^*,

*C*1. *Every* $\mathbf{P} \in \mathscr{H}$ *is bounded. That is, given* $\mathbf{P} \in \mathscr{H}$, *there are real numbers* a *and* b *such that* $P^*(\{s : a \leq E(u, \mathbf{P}(s)) \leq b\}) = 1$;

*C*2. *For all* $A \subseteq S$, $P^*(A) = 0 \Leftrightarrow A$ *is null;*

*C*3. u *is bounded if there is a denumerable partition of* S *such that* $P^*(A) > 0$ *for every event in the partition;*

*C*4. *A real-valued function* u' *on* X *and a probability measure* Q^* *on* S *satisfy* (13.3) *in place of* u *and* P^* *if and only if* $Q^* = P^*$ *and* u' *is a positive linear transformation of* u.

*B*4 says that there is some pair of constant horse lotteries that are not indifferent. *B*4 and $X = X(s)$ for all $s \in S$ supplant *A*4 of Theorem 13.2.

*B*5 is an obvious monotonicity axiom for nonnull events. *B*6 is a form of sure-thing or dominance axiom. It is similar to axiom *A*4*a* in Section 10.4 and to *P*7 in the next chapter. *B*6 is needed only if S is infinite.

In addition to the noted conclusions of Theorem 13.3 it should be remarked that P^* has no special properties apart from those of a probability measure. If S is infinite, it may or may not be true that $P^*(A) = 1$ for some finite $A \subseteq S$. If $P^*(A) < 1$ for every finite $A \subseteq S$ it may or may not be true that S can be partitioned into an arbitrary finite number of events each with equal probability. In addition, u has no special properties other than those noted

in $C3$ and $C4$, except that it cannot be constant on X. u might be unbounded when the condition of $C3$ does not hold.

Proof of Theorem 13.3

To prove the theorem we shall prove a series of statements that, taken together, establish all conclusions of the theorem. For convenience we first list these statements. u and P^* for $S2$–$S5$ are defined as in $S1$.

$S1$. $B1$–$B5 \Rightarrow$ (13.3) *for all* $\mathbf{P}, \mathbf{Q} \in \mathcal{H}_0$ *where* $\mathcal{H}_0 = \{\mathbf{P}:\mathbf{P} \in \mathcal{H}$ *and* $\mathbf{P}$ *is constant on each event in some finite partition of* $S\}$. *u and* P^* *for* (13.3) *on* $\mathcal{H}_0$ *are unique up to a positive linear transformation and unique, respectively, and* $P^*(A) = 0$ *if and only if A is null.*

$S2$. $B1$–$B6 \Rightarrow$ (13.3) *for all bounded horse lotteries.* (See $C1$.)

$S3$. ($B1$–$B6$, *there is a denumerable partition of S such that* $P^*(A) > 0$ *for every A in the partition*) $\Rightarrow$ *u on X is bounded.*

$S4$. *If for each positive integer n there is an n-event partition of S for which each event has positive probability under* P^*, *then there is a denumerable partition of S such that* $P^*(A) > 0$ *for every A in the partition.*

$S5$. *If the hypotheses of* $S4$ *are false then* $B1$–$B6$ *imply that all horse lotteries are bounded.* (In this case it is not necessarily true that P^* is a simple probability measure. See Exercises 5 and 6.)

Note that $S3$, $S4$, and $S5$ imply that all horse lotteries are bounded. If the hypotheses of $S4$ are true then, by $S3$, u on X is bounded and hence all $\mathbf{P} \in \mathcal{H}$ must be bounded. On the other hand, if the hypotheses of $S4$ are false then, by $S5$, all $\mathbf{P} \in \mathcal{H}$ are bounded even though u on X might be unbounded.

Proof of S1. Let $\{B_1, \ldots, B_n\}$ be a finite partition of S. Then, by essentially the same proof used for Theorem 13.2, there are nonnegative numbers $P_B^*(B_1), \ldots, P_B^*(B_n)$ that sum to one and there is a real-valued function u_B on X such that, whenever $\mathbf{P} = P_i$ on $\mathbf{Q} = Q_i$ on B_i $(i = 1, \ldots, n)$

$$\mathbf{P} < \mathbf{Q} \Leftrightarrow \sum_{i=1}^{n} P_B^*(B_i)E(u_B, P_i) < \sum_{i=1}^{n} P_B^*(B_i)E(u_B, Q_i), \tag{13.6}$$

and when this holds $P_B^*(B_i) = 0$ if and only if B_i is null, P_B^* is unique, and u_B is unique up to a positive linear transformation.

Let $\mathcal{H}_c$ be the set of all constant horse lotteries in $\mathcal{H}$. Thus $\mathcal{H}_c \subseteq \mathcal{H}_0$. If $P, Q \in \mathcal{P}$ and if $\{B_1, \ldots, B_n\}$ and $\{C_1, \ldots, C_m\}$ are partitions of S then, by (13.6), $E(u_B, P) < E(u_B, Q) \Leftrightarrow E(u_C, P) < E(u_C, Q)$. Hence, noting that $\mathcal{H}_c$ is a mixture set for which $B1$, $B2$, and $B3$ hold, it follows from Theorem 8.4 that u_C on X is a positive linear transformation of u_B on X. Therefore, we can drop the partition-specific subscript on u and have, in place of (13.6),

$$\mathbf{P} < \mathbf{Q} \Leftrightarrow \sum_{i=1}^{n} P_B^*(B_i)E(u, P_i) < \sum_{i=1}^{n} P_B^*(B_i)E(u, Q_i), \tag{13.7}$$

in which u is unique up to a positive linear transformation.

For event $A \subseteq S$ let $\mathcal{H}_A = \{\mathbf{P}:\mathbf{P} \in \mathcal{H}$ and $\mathbf{P}$ is constant on A and on $A^c\}$. If $\{B_1, \ldots, B_n\}$ and $\{C_1, \ldots, C_m\}$ are partitions of S and if A is an element in each partition then (13.7) implies that, for all $\mathbf{P}, \mathbf{Q} \in \mathcal{H}_A$, with $\mathbf{P} = P_A$ and $\mathbf{Q} = Q_A$ on A, and $\mathbf{P} = P_A^c$ and $\mathbf{Q} = Q_A^c$ on A^c,

$$\begin{aligned} P_B^*(A)E(u, P_A) + P_B^*(A^c)E(u, P_A^c) &< P_B^*(A)E(u, Q_A) + P_B^*(A^c)E(u, Q_A^c) \\ &\Leftrightarrow P_C^*(A)E(u, P_A) + P_C^*(A^c)E(u, P_A^c) \\ &< P_C^*(A)E(u, Q_A) + P_C^*(A^c)E(u, Q_A^c). \end{aligned}$$

It then follows from the version of (13.7) for the partition $\{A, A^c\}$ that $P_B^*(A) = P_C^*(A)$, so that we can drop the partition-specific subscript on P^* and write (13.7) as

$$\mathbf{P} < \mathbf{Q} \Leftrightarrow \sum_{i=1}^{n} P^*(B_i)E(u, P_i) < \sum_{i=1}^{n} P^*(B_i)E(u, Q_i). \tag{13.8}$$

Adding $P^*(\varnothing) = 0$ to complete P^*, it follows that P^* is uniquely determined and that $P^*(A) = 0$ if and only if A is null. Finite additivity for P^* is easily demonstrated using partitions $\{A, B, (A \cup B)^c\}$ and $\{A \cup B, (A \cup B)^c\}$ with $A \cap B = \varnothing$ in an analysis like that leading to (13.8).

Finally, to obtain (13.3) for all of $\mathcal{H}_0$, let $\mathbf{P} = P_i$ on B_i and $\mathbf{Q} = Q_j$ on C_j for partitions $\{B_1, \ldots, B_n\}$ and $\{C_1, \ldots, C_m\}$. Applying (13.8) to the partition $\{B_i \cap C_j : i = 1, \ldots, n; j = 1, \ldots, m; B_i \cap C_j \neq \varnothing\}$, we get $\mathbf{P} < \mathbf{Q} \Leftrightarrow \sum_i \sum_j P^*(B_i \cap C_j)E(u, P_i) < \sum_i \sum_j P^*(B_i \cap C_j)E(u, Q_j)$. By finite additivity, the last expression is $\sum_i P^*(B_i)E(u, P_i) < \sum_j P^*(C_j)E(u, Q_j)$. ◆

Proof of S2. Since $\mathcal{H}$ is a mixture set, Theorem 8.4 implies that there is a real-valued function v on $\mathcal{H}$ that satisfies $\mathbf{P} < \mathbf{Q} \Leftrightarrow v(\mathbf{P}) < v(\mathbf{Q})$ and $v(\alpha\mathbf{P} + (1 - \alpha)\mathbf{Q}) = \alpha v(\mathbf{P}) + (1 - \alpha)v(\mathbf{Q})$. Since these expressions hold on $\mathcal{H}_0 \subseteq \mathcal{H}$ it follows from (13.3) for $\mathcal{H}_0$ and Theorem 8.4 that w on $\mathcal{H}_0$ defined by $w(\mathbf{P}) = E[E(u, \mathbf{P}(s)), P^*]$, is a positive linear transformation of the restriction of v on $\mathcal{H}_0$. Without loss in generality we can therefore specify that

$$v(\mathbf{P}) = E[E(u, \mathbf{P}(s)), P^*] \tag{13.9}$$

for all $\mathbf{P} \in \mathcal{H}_0$, with

$$\mathbf{P} < \mathbf{Q} \Leftrightarrow v(\mathbf{P}) < v(\mathbf{Q}), \qquad \text{for all } \mathbf{P}, \mathbf{Q} \in \mathcal{H} \tag{13.10}$$

$$v(\alpha\mathbf{P} + (1 - \alpha)\mathbf{Q}) = \alpha v(\mathbf{P}) + (1 - \alpha)v(\mathbf{Q}), \qquad \text{for all } (\alpha, \mathbf{P}, \mathbf{Q}) \in [0, 1] \times \mathcal{H}^2. \tag{13.11}$$

According to (13.10), $S2$ is proved if (13.9) holds for all bounded $\mathbf{P} \in \mathcal{H}$.

Our first step toward this end will be to show that

$$c = \inf\{E(u, \mathbf{P}(s)) : s \in A\} \leq v(\mathbf{P}) \leq \sup\{E(u, \mathbf{P}(s)) : s \in A\} = d \tag{13.12}$$

holds when $P^*(A) = 1$ and c and d as defined are finite. Let $\mathbf{Q} = \mathbf{P}$ on A and $c \leq E(u, \mathbf{Q}(s)) \leq d$ on A^c. Since A^c is null, $\mathbf{Q} \sim \mathbf{P}$ and $v(\mathbf{P}) = v(\mathbf{Q})$ by (13.10). To verify $c \leq v(\mathbf{Q}) \leq d$ with c and d finite, suppose to the contrary that $d < v(\mathbf{Q})$. With $c \leq E(u, Q') \leq d$ and $\mathbf{Q}' = Q'$ on S, let $\mathbf{R} = \alpha\mathbf{Q} + (1 - \alpha)\mathbf{Q}'$ with $\alpha < 1$ near enough to one so that $d < v(\mathbf{R}) = \alpha v(\mathbf{Q}) + (1 - \alpha)v(\mathbf{Q}') < v(\mathbf{Q})$. Then $\mathbf{R} < \mathbf{Q}$ by (13.10). But since $E(u, \mathbf{Q}(s)) \leq d < v(\mathbf{R})$, it follows from (13.10) that $\mathbf{Q}(s) < \mathbf{R}$ for all $s \in S$ and hence by *B*6 that $\mathbf{Q} \preccurlyeq \mathbf{R}$, a contradiction. Hence $v(\mathbf{Q}) \leq d$. By a symmetric proof, $c \leq v(\mathbf{Q})$.

With $\mathbf{P}$ bounded let A with $P^*(A) = 1$ be an event on which $E(u, \mathbf{P}(s))$ is bounded and define c and d as in (13.12). If $c = d$ then (13.9) is immediate. Henceforth assume that $c < d$: for convenience we shall take $c = 0$, $d = 1$. Let $\mathbf{Q}$ be defined as in the preceding paragraph so that $v(\mathbf{P}) = v(\mathbf{Q})$ and $E[E(u, \mathbf{P}(s)), P^*] = E[E(u, \mathbf{Q}(s)), P^*]$, the latter by Exercise 10.22. To show that $v(\mathbf{Q}) = E[E(u, \mathbf{Q}(s)), P^*]$, let $\{A_1, \ldots, A_n\}$ be the partition (ignoring empty sets) of S defined by

$$A_1 = \{s: 0 \leq E(u, \mathbf{Q}(s)) \leq 1/n\}$$
$$A_i = \{s: (i - 1)/n < E(u, \mathbf{Q}(s)) \leq i/n\} \qquad i = 2, \ldots, n,$$

and let $P_i \in \mathcal{P}$ be such that

$$(i - 1)/n \leq E(u, P_i) \leq i/n \qquad \text{for } i = 1, \ldots, n. \tag{13.13}$$

The existence of such P_i is guaranteed by (13.12). Let $\mathbf{P}_i = \mathbf{Q}$ on A_i and $\mathbf{P}_i = P_i$ on A_i^c $(i = 1, \ldots, n)$; let $\mathbf{P}_0 = \sum_{i=1}^n (1/n)\mathbf{P}_i$; and let $\mathbf{R} = \sum_{j \neq i} (1/(n - 1))P_j$ on A_i for $i = 1, \ldots, n$. Then $\mathbf{P}_0(s) = \sum_i (1/n)\mathbf{P}_i(s) = (1/n)\mathbf{Q}(s) + ((n - 1)/n) \sum_{j \neq i} (1/(n - 1))P_j$ when $s \in A_i$, so that $\mathbf{P}_0 = (1/n)\mathbf{Q} + ((n - 1)/n)\mathbf{R}$. Hence, by (13.11) and $\mathbf{P}_0 = \sum_i (1/n)\mathbf{P}_i$,

$$v(\mathbf{Q}) = \sum_{i=1}^{n'} v(\mathbf{P}_i) - (n - 1)v(\mathbf{R}). \tag{13.14}$$

Since $\mathbf{R} \in \mathcal{K}_0$, (13.9) implies that $v(\mathbf{R}) = \sum_i E(u, \sum_{j \neq i} (1/(n - 1))P_j)P^*(A_i) = (1/(n - 1)) \sum_i [\sum_{j \neq i} E(u, P_j)]P^*(A_i)$. Substituting this in (13.14) gives

$$v(\mathbf{Q}) = \sum_{i=1}^{n} v(\mathbf{P}_i) - \sum_{i=1}^{n} \sum_{j \neq i} E(u, P_j)P^*(A_i). \tag{13.15}$$

Now by (13.13), (13.12), and the definition of $\mathbf{P}_i$

$$(i - 1)/n \leq v(\mathbf{P}_i) \leq i/n \qquad \text{for } i = 1, \ldots, n. \tag{13.16}$$

Since $0 = \inf \{E(u, \mathbf{Q}(s)): s \in S\}$ and $1 = \sup \{E(u, \mathbf{Q}(s)): s \in S\}$, P_i that satisfy (13.13) can be selected so that either

$$E(u, P_1) = 1/n, \quad \text{and} \quad E(u, P_i) = (i - 1)/n \qquad \text{for } i > 1 \tag{13.17}$$

or

$$E(u, P_i) = i/n \qquad \text{for } i < n, \quad \text{and} \quad E(u, P_n) = (n - 1)/n. \tag{13.18}$$

Applying (13.17) and the left side of (13.16) to (13.15), we get

$$\begin{aligned} v(\mathbf{Q}) &\geq \sum_{i=1}^{n} \frac{i-1}{n} - \frac{n-1}{2} P^*(A_1) - \sum_{i=2}^{n} \left[\frac{n-1}{2} - \frac{i-1}{n} + \frac{1}{n}\right] P^*(A_i) \\ &= \frac{n-1}{2} - \frac{n-1}{2} + \sum_{i=2}^{n} \frac{i-1}{n} P^*(A_i) - \frac{1}{n}[1 - P^*(A_1)] \\ &\geq \sum_{i=1}^{n} \frac{i-1}{n} P^*(A_i) - \frac{1}{n}. \end{aligned}$$

Applying (13.18) and the right side of (13.16) to (13.15), we get

$$\begin{aligned} v(\mathbf{Q}) &\leq \sum_{i=1}^{n} \frac{i}{n} - \sum_{i=1}^{n-1} \left[\frac{n-1}{2} - \frac{i}{n} + \frac{n-1}{n}\right] P^*(A_i) - \frac{n-1}{2} P^*(A_n) \\ &= \sum_{i=1}^{n} \frac{i}{n} P^*(A_i) + \frac{1}{n}[1 - P^*(A_n)] \\ &\leq \sum_{i=1}^{n} \frac{i}{n} P^*(A_i) + \frac{1}{n}. \end{aligned}$$

By the definition of E, $\sum((i-1)/n)P^*(A_i) \leq E[E(u, \mathbf{Q}(s)), P^*] \leq \sum(i/n)P^*(A_i)$, so that $|v(\mathbf{Q}) - E[E(u, \mathbf{Q}(s)), P^*]| \leq 2/n$ for $n = 1, 2, \ldots$. Hence, $v(\mathbf{Q}) = E[E(u, \mathbf{Q}(s)), P^*]$. ◆

Proof of S3. Let $\mathcal{A}$ be a denumerable partition of S with $P^*(A) > 0$ for all $A \in \mathcal{A}$. $\{P^*(A) : A \in \mathcal{A}\}$ must have a largest element, say $P^*(A_1)$. Then $\{P^*(A) : A \in \mathcal{A} - \{A_1\}\}$ must have a largest element, say A_2. Continuing this, we get a sequence $A_1, A_2, \ldots$ with $\{A_1, A_2, \ldots\} = \mathcal{A}$ and $P^*(A_i) \geq P^*(A_{i+1})$ for $i = 1, 2, \ldots$.

Contrary to $S3$ suppose that u is unbounded above. By a linear transformation of u we can assume that $[0, \infty) \subseteq \{E(u, P) : P \in \mathcal{P}\}$. Let $P_i \in \mathcal{P}$ be such that

$$E(u, P_i) = 1/P^*(A_i) \qquad \text{for } i = 1, 2, \ldots. \tag{13.19}$$

Let $\mathbf{P} = P_i$ on A_i $(i = 1, 2, \ldots)$ and let $\mathbf{Q}_n$ be constant on each A_i for $i \leq n$ and constant on $\bigcup_{i=n+1}^{\infty} A_i$ with

$$\begin{aligned} E(u, \mathbf{Q}_n(s)) &= P^*(A_n)^{-1} - P^*(A_i)^{-1} \qquad \text{for all } s \in A_i; \quad i = 1, \ldots, n \\ E(u, \mathbf{Q}_n(s)) &= 0 \qquad \text{for all } s \in \bigcup_{i=n+1}^{\infty} A_i. \end{aligned} \tag{13.20}$$

Let v on $\mathcal{H}$ satisfy (13.10) and (13.11) and also (13.9) on $\mathcal{H}_0$. Then

$$\begin{aligned} v(\mathbf{Q}_n) &= \sum_{i=1}^{n} [P^*(A_n)^{-1} - P^*(A_i)^{-1}] P^*(A_i) \\ &= P^*(A_n)^{-1} \sum_{i=1}^{n} P^*(A_i) - n \qquad \text{for } n = 1, 2, \ldots. \end{aligned} \tag{13.21}$$

By (13.19) and (13.20), $E(u, \frac{1}{2}\mathbf{P}(s) + \frac{1}{2}\mathbf{Q}_n(s)) = \frac{1}{2}P^*(A_i)^{-1} + \frac{1}{2}[P^*(A_n)^{-1} - P^*(A_i)^{-1}] = \frac{1}{2}P^*(A_n)^{-1}$ for all $s \in \bigcup_1^n A_i$ and, by $P^*(A_i) \geq P^*(A_{i+1})$ and (13.20), $E(u, \frac{1}{2}\mathbf{P}(s) + \frac{1}{2}\mathbf{Q}_n(s)) \geq \frac{1}{2}P^*(A_n)^{-1}$ for all $s \in \bigcup_{n+1}^\infty A_i$. Therefore $\inf\{E(u, \frac{1}{2}\mathbf{P}(s) + \frac{1}{2}\mathbf{Q}_n(s)) : s \in S\} = \frac{1}{2}P^*(A_n)^{-1}$. Hence, by (13.12), $v(\frac{1}{2}\mathbf{P} + \frac{1}{2}\mathbf{Q}_n) \geq \frac{1}{2}P^*(A_n)^{-1}$, which on using (13.11) and (13.21) implies that

$$v(\mathbf{P}) \geq P^*(A_n)^{-1} - P^*(A_n)^{-1}\sum_{i=1}^n P^*(A_i) + n \geq n \qquad \text{for } n = 1, 2, \ldots .$$

But this requires $v(\mathbf{P})$ to be infinite. Hence u is bounded above. A symmetric proof shows that u is bounded below under $S3$'s hypotheses. ◆

Proof of S4. For each integer $n \geq 2$ let $\mathcal{A}^n$ be an n-part partition of S each event in which has positive probability. Define a new set of partitions $\mathcal{B}^2, \mathcal{B}^3, \ldots$ recursively as follows:

$$\mathcal{B}^2 = \mathcal{A}^2$$

$$\mathcal{B}^n = \{A \cap B : A \in \mathcal{A}^n, B \in \mathcal{B}^{n-1}, A \cap B \neq \varnothing\}, \qquad n = 3, 4, \ldots .$$

It is easily verified that $\mathcal{B}^n$ contains at least n events with positive probability and that $\mathcal{B}^{n+1}$ is as fine as $\mathcal{B}^n$ so that $B \in \mathcal{B}^{n+1} \Rightarrow C \in \mathcal{B}^n$ for some C that includes B. For each $A \in \mathcal{B}^2$ let

$N_n^1(A)$ = number of events in $\mathcal{B}^n$ $(n > 2)$ that are included in A and have positive probability.

With $\mathcal{B}^2 = \{A, A^c\}$ it follows that $N_n^1(A) + N_n^1(A^c) \geq n$ for $n = 3, 4, \ldots$. Thus, as n gets large, at least one of $N_n^1(A)$ and $N_n^1(A^c)$ approaches infinity. Let A_1 be an event in $\mathcal{B}^2$ for which $N_n^1(A_1) \to \infty$ and let $B_1 = A_1^c$. Then $P^*(B_1) > 0$ and B_1 will be the first element in our desired denumerable partition.

Let $n(1)$ be an integer for which $\mathcal{B}^{n(1)}$ contains more than one subset of A_1 that has positive probability. For each $A \subseteq A_1$ and $A \in \mathcal{B}^{n(1)}$ let

$N_n^2(A)$ = number of events in $\mathcal{B}^n$ $(n > n(1))$ that are included in A and have positive probability.

Let $\mathcal{A} = \{A : A \subseteq A_1, A \in \mathcal{B}^{n(1)}\}$. Then $\sum_{\mathcal{A}} N_n^2(A) = N_n^1(A_1)$ so that, since $N_n^1(A_1) \to \infty$ as $n \to \infty$, at least one $N_n^2(A) \to \infty$ as $n \to \infty$. Let $A_2 \in \mathcal{A}$ be such an A and let $B_2 = A_1 \cap A_2^c$. Then $P^*(B_2) > 0$ and $\{B_1, B_2, A_2\}$ is a partition of S with $N_n^2(A_2) \to \infty$ as $n \to \infty$.

Continuing this construction in the obvious way gives a denumerable sequence $B_1, B_2, B_3, \ldots$ of mutually disjoint events each of which has positive probability. The conclusion of $S4$ follows. ◆

Proof of S5. If the hypotheses of $S4$ are false then there is a unique positive integer m for which there is an m-event partition of S that has

positive probability for each event and such that no partition of S has positive probability on more than m of its events.

For convenience assume that $u(y) = 0$ for some $y \in X$. Suppose then, contrary to the conclusion of $S5$, that $\mathbf{Q}$ is unbounded above. Let $\mathbf{P}$ be obtained from $\mathbf{Q}$ by replacing each x for which $\mathbf{Q}(s)(x) > 0$ and $u(x) < 0$ by y with $u(y) = 0$, for all $s \in S$. Then $E(u, \mathbf{P}(s)) \geq 0$ for all s and $\mathbf{P}$ is unbounded above. Then, for every $n > 0$,

$$P^*\{E(u, \mathbf{P}(s)) \geq n\} = P^*(\{s : E(u, \mathbf{P}(s)) \geq n\}) > 0.$$

By the preceding paragraph, $P^*\{E(u, \mathbf{P}(s)) \geq n\}$ can change no more than m times as n increases. Hence there is an N and an $\alpha > 0$ such that

$$P^*\{E(u, \mathbf{P}(s)) \geq n\} = \alpha \qquad \text{for all } n \geq N. \tag{13.22}$$

Let $E(u, P_i) = i$ for $i = 1, 2, \ldots,$ let $\mathbf{Q}_n = \mathbf{P}$ on $\{s : E(u, \mathbf{P}(s)) \geq n\}$, $\mathbf{Q}_n = P_n$ on $\{s : E(u, \mathbf{P}(s)) < n\}$, and let $\mathbf{R}_n = P_n$ on $\{s : E(u, \mathbf{P}(s)) \geq n\}$, $\mathbf{R}_n = \mathbf{P}$ on $\{s : E(u, \mathbf{P}(s)) < n\}$. Then, with $\mathbf{P}_n = P_n$ on S, $\frac{1}{2}\mathbf{P} + \frac{1}{2}\mathbf{P}_n = \frac{1}{2}\mathbf{Q}_n + \frac{1}{2}\mathbf{R}_n$, so that with v on $\mathcal{H}$ as given by (13.10) and (13.11) and satisfying (13.9) for all bounded horse lotteries,

$$v(\mathbf{P}) + n = v(\mathbf{Q}_n) + v(\mathbf{R}_n) \qquad n = 1, 2, \ldots . \tag{13.23}$$

Since $\mathbf{R}_n$ is bounded, (13.9) and (13.22) give $v(\mathbf{R}_n) = E[E(u, \mathbf{R}_n(s)), P^*] \geq n\alpha$ for all $n \geq N$. Since $P_{n-1} < Q_n(s)$ for all $s \in S$, $B6$ implies that $P_{n-1} \preccurlyeq \mathbf{Q}_n$ so that $v(\mathbf{Q}_n) \geq n - 1$ for all n. Then, using (13.23),

$$v(\mathbf{P}) \geq n\alpha - 1 \qquad \text{for all } n \geq N,$$

which contradicts the finiteness of $v(\mathbf{P})$. Hence $\mathbf{Q}$ is not unbounded above. A symmetric proof shows that $\mathbf{Q}$ is bounded below. ◆

13.4 THE PART II DECISION MODEL

Beginning with the set F of acts and the set X of consequences as in the Part II approach, let S be the set of all functions on F to X (see S' in Section 12.1). Then the subset of X that is immediately relevant under "state" $s \in S$ is $X(s) = \{s(f) : f \in F\}$. For many $s \in S$, $X(s)$ will be a proper subset of X, and for each constant s that assigns the same x to each f, $X(s) = \{x\}$. Hence the horse-lottery theory of Sections 13.2 and 13.3 cannot be used in establishing the Part II model unless we assume that consequences other than those in $X(s)$ can be considered relevant under state s.

Suppose in fact that we assume the extreme, that all consequences are relevant under every state. Then, under $B1$–$B6$ of Theorem 13.3, (13.3) follows. With $f \in F$, $Y \subseteq X$, and $P_f(Y) = P^*(\{s : s(f) \in Y\})$ we then obtain $E(u, P_f) = E[E(u, \mathbf{P}(s)), P^*]$ when $\mathbf{P}(s)(s(f)) = 1$ for each $s \in S$. Then under

the natural definition of $<$ on F in terms of $<$ on $\mathcal{K}$,

$$f < g \Leftrightarrow E(u, P_f) < E(u, P_g), \qquad \text{for all } f, g \in F, \tag{13.24}$$

which is the Part II model. Although extraneous probabilities are used in deriving this, note that $P_f, P_g, \ldots$ are defined from P^*, which may have nothing to do with the extraneous probabilities and is itself derived from the axioms.

It may of course be stretching things too far to assume that consequences not in $X(s)$ are relevant under s. However, it may be possible to say something about state probabilities even when this assumption is not made.

Suppose, for example, that $F = \{f, g\}$ and $X = \{\text{win, lose}\}$. Then S has four elements: $s_1(f) = s_1(g) = \text{win}$; $s_2(f) = \text{win}$, $s_2(g) = \text{lose}$; $s_3(f) = \text{lose}$, $s_3(g) = \text{win}$; $s_4(f) = s_4(g) = \text{lose}$. Let $\mathcal{K} = \mathcal{P}(s_1) \times \mathcal{P}(s_2) \times \mathcal{P}(s_3) \times \mathcal{P}(s_4)$. The conditions of Theorem 13.1 then give $\mathbf{P} < \mathbf{Q} \Leftrightarrow \sum_i E(u_i, \mathbf{P}(s_i)) < \sum_i E(u_i, \mathbf{Q}(s_i))$. Since $X(s_1) = \{\text{win}\}$ and $X(s_4) = \{\text{lose}\}$, the first and fourth terms drop out of this and we are left with

$$\mathbf{P} < \mathbf{Q} \Leftrightarrow E(u_2, \mathbf{P}(s_2)) + E(u_3, \mathbf{P}(s_3)) < E(u_2, \mathbf{Q}(s_2)) + E(u_3, \mathbf{Q}(s_3)). \tag{13.25}$$

According to our definition s_1 and s_4 are null, but this is only because $X(s_1)$ and $X(s_4)$ each contain a single consequence: the decision maker might consider s_1 to be the most probable state. In such a case we would regard $P^*(s_1)$ and $P^*(s_4)$ as indeterminate within the structure of our axioms. This indeterminacy actually causes no difficulty since the $i = 1$ and $i = 4$ terms do not appear in (13.25).

With regard to s_2 and s_3, $\mathcal{P}(s_2) = \mathcal{P}(s_3)$ since $X(s_2) = X(s_3) = \{\text{win, lose}\}$. If condition $A5$ of Theorem 13.2 is used it follows from (13.25) that (assuming some strict preference) there are $\lambda_2 \geq 0$ and $\lambda_3 \geq 0$ with $\lambda_2 + \lambda_3 > 0$ and there is a real-valued function u on X such that $\mathbf{P} < \mathbf{Q} \Leftrightarrow \lambda_2 E(u, \mathbf{P}(s_2)) + \lambda_3 E(u, \mathbf{P}(s_3)) < \lambda_2 E(u, \mathbf{Q}(s_2)) + \lambda_3 E(u, \mathbf{Q}(s_3))$. Here we would interpret λ_2 and λ_3 as proportional to $P^*(s_2)$ and $P^*(s_3)$ respectively so that, if $\lambda_3 > 0$, $\lambda_2/\lambda_3 = P^*(s_2)/P^*(s_3)$. If $u(\text{win}) > u(\text{lose})$ it is easily seen from this that $f < g \Leftrightarrow P^*(s_2) < P^*(s_3)$.

13.5 SUMMARY

The usual states expected-utility decision model can be derived from axioms that involve extraneous probabilities when there is sufficient overlap among consequences considered relevant under different states. When the number of states is finite, the assumption that there are two nonindifferent consequences that are relevant under every state is sufficient. For the more general case, in which the size of S is arbitrary, it was assumed that all consequences are

relevant under every state. Even when there may be no overlap among the consequences under different states, the expected-utility axioms of Chapter 8 when applied to horse lotteries with S finite lead to an additive-utility representation that is similar to additive forms of Section 11.1 and (12.11).

Although the horse-lottery approach presumes a continuum of extraneous probabilities, this appears to be offset by its general applicability since it places almost no restrictions on the sizes of S and X. Moreover, it places no unusual restrictions on the utility function on X or on the probability measure P^* on S.

INDEX TO EXERCISES

1. Insufficiently connected $X(s)$. 2. $P^* = 1$ for a finite subset. 3. Intersections of partitions. 4. $S4$. 5–6. Zero-one measures and $S4$. 7. Additivity when $X = \Pi X_i$. 8–15. $\mathbf{P}(s) \preccurlyeq \mathbf{Q}(s)$ for all $s \in S \Rightarrow \mathbf{P} \preccurlyeq \mathbf{Q}$. 16–19. Even-chance theory.

Exercises

1. Let $S = \{s_1, s_2, s_3\}$, $X(s_1) = \{x, y, z, w\}$, $X(s_2) = \{x, y, r, t\}$, $X(s_3) = \{z, w, r, t\}$. Let the hypotheses of Theorem 13.2 with the exception of $A4$ hold, and let the following values of u_i ($i = 1, 2, 3$) satisfy (13.2):

	x	y	z	w	r	t
u_1	0	1	2	3		
u_2	0	1			2	3
u_3			2	3	4	6

a. Verify that, for each i, j there is a unique r_{ij} such that $u_i(p) = r_{ij}u_j(p)$ for all $p \in X(s_i) \bigcap X(s_j)$. Is $r_{23} = r_{21}/r_{31}$?

b. Show that it is impossible to define u on $X = \bigcup X(s_i)$ and P^* on S so as to satisfy (13.3).

2. Prove that if $P^*(A) = 1$ for some finite $A \subseteq S$ then $B1$–$B5$ imply (13.3) in the structural context of Section 13.3.

3. Let $\mathfrak{D}$ be a set of partitions of S. Show that $\{\bigcap_{D \in \mathfrak{D}} f(D): f(D) \in D$ for each $D \in \mathfrak{D}$, $\bigcap_{D \in \mathfrak{D}} f(D) \neq \varnothing\}$ is a partition of S.

4. In connection with $S4$ and its proof, suppose that $\mathfrak{B}^2, \mathfrak{B}^3, \ldots$ is a sequence of partitions of S such that (1) $\mathfrak{B}^n$ contains exactly n events, each with positive probability, and (2) $A \in \mathfrak{B}^{n+1} \Rightarrow A \subseteq B$ for some $B \in \mathfrak{B}^n$. Show that it may be impossible to select one event from each $\mathfrak{B}^n$ so that the selected events are mutually disjoint.

5. Let P^* on S be defined in such a way that there is a set $\mathcal{A}$ of subsets of S such that $P^*(A) = 1$ if $A \in \mathcal{A}$, and $P^*(A) = 0$ if $A \notin \mathcal{A}$. Prove that if $\bigcap_{\mathcal{A}} A \neq \varnothing$ then this intersection contains exactly one s and, for this s, $P^*(s) = 1$.

6. Let S be infinite and let $\mathcal{S}$ be the set of all sets $\mathcal{A}$ of subsets of S that have the following four properties:

1. $\varnothing \in \mathcal{A}$ and $\{s\} \in \mathcal{A}$ for all $s \in S$;
2. $A \in \mathcal{A} \Rightarrow A^c \notin \mathcal{A}$;
3. $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$;
4. $(A \in \mathcal{A}, B \subseteq A) \Rightarrow B \in \mathcal{A}$.

a. Is the set of all finite subsets of S in $\mathcal{S}$?

b. Use (4) to show that $(\mathcal{A} \in \mathcal{S}, A \cup B \in \mathcal{A}) \Rightarrow A, B \in \mathcal{A}$.

c. Prove that $(\mathcal{A} \in \mathcal{S}, A \in \mathcal{A}, B \notin \mathcal{A}) \Rightarrow A \cup B \notin \mathcal{A}$.

d. Use Zorn's Lemma to prove that there is an $\mathcal{A} \in \mathcal{S}$ that is maximal with respect to (1) through (4). Let $\mathcal{A}^*$ be maximal (if $\mathcal{A}^* \subset \mathcal{A}'$ then $\mathcal{A}' \notin \mathcal{S}$) and let $\mathcal{B}^*$ be the set of all subsets of S that are not in $\mathcal{A}^*$.

e. Prove that $A, B \in \mathcal{B}^* \Rightarrow A \cap B \neq \varnothing$. For this suppose that $A, B \in \mathcal{B}^*$ and $A \cap B = \varnothing$. Then show that $\mathcal{A}^0 = \mathcal{A}^* \cup \{C \cup D : C \subseteq A, D \in \mathcal{A}^*\}$ is in $\mathcal{S}$, contradicting the maximality of $\mathcal{A}^*$.

f. Let $P^*(A) = 0$ if $A \in \mathcal{A}^*$ and $P^*(A) = 1$ if $A \in \mathcal{B}^*$. Show that P^* is a probability measure on S. Note that $\bigcup_{\mathcal{B}^*} A = \varnothing$ and compare with the preceding exercise.

g. Explain why the failure of the hypotheses of $S4$ does not imply that P^* on S is a simple probability measure.

7. Suppose the hypotheses of Theorem 13.3 hold and, in addition $X = \prod_{j=1}^{n} X_j$ and, with P_j the marginal of $P \in \mathcal{P}$ on X_j, $(\mathbf{P} = P$ on S, $\mathbf{Q} = Q$ on S, $P_j = Q_j$ for $j = 1, \ldots, n) \Rightarrow \mathbf{P} \sim \mathbf{Q}$. Show that there are real-valued functions $u_1, \ldots, u_n$ on $X_1, \ldots, X_n$ respectively such that

$$\mathbf{P} \prec \mathbf{Q} \Leftrightarrow \sum_{j=1}^{n} E[E(u_j,\ \mathbf{P}(s)_j),\ P^*] < \sum_{j=1}^{n} E[E(u_j,\ \mathbf{Q}(s)_j),\ P^*], \text{ for all } \mathbf{P}, \mathbf{Q} \in \mathcal{K}$$

where $\mathbf{P}(s)_j$ is the marginal of $\mathbf{P}(s)$ on X_j.

Note: Exercises 8–15 *are set in the context of Section* 13.3. *Axiom B7 is*: $\mathbf{P}(s) \preccurlyeq \mathbf{Q}(s)$ *for all* $s \in S \Rightarrow \mathbf{P} \preccurlyeq \mathbf{Q}$.

8. Prove that ($B1$, $B7$, A^c is null, $\mathbf{P}(s) \preccurlyeq \mathbf{Q}(s)$ for all $s \in A) \Rightarrow \mathbf{P} \preccurlyeq \mathbf{Q}$.

9. Prove that $(B1, B7) \Rightarrow$ if $\mathbf{P} = P$ and $\mathbf{Q} = Q$ on A, $\mathbf{P} = \mathbf{Q}$ on A^c, and $\mathbf{P} \prec \mathbf{Q}$ then $P \prec Q$. (This is one half of $B5$.)

10. By a straightforward partition proof show that if X has a least preferred and a most preferred consequence then $(B1\text{–}B5, B7) \Rightarrow$ (13.3).

11. Show that ($B1$–$B5$, $B7$, there is a denumerable partition $\mathcal{A}$ of S for which $P^*(A) > 0$ for every $A \in \mathcal{A}) \Rightarrow u$ on X is bounded. (Use $S1$.)

12. (*Continuation.*) Prove $S2$ when $B6$ in its hypotheses is replaced by $B7$. To do this you need only verify (13.12) when $P^*(A) = 1$ and c and d are finite. This is the critical point for $B6$ and therefore for $B7$.

13. (*Continuation.*) Use Exercises 11 and 12 to argue that the conclusions of Theorem 13.3 are valid when $B6$ in its hypotheses is replaced by $B7$.

14. Let $S = \{1, 2, 3, \ldots\}$, $X = [0, 1)$, $u(x) = x$, and let P^* be a probability measure on S that has $P^*(s) = 0$ for $s = 1, 2, \ldots$. Suppose $\mathbf{P} \prec \mathbf{Q} \Leftrightarrow v(\mathbf{P}) < v(\mathbf{Q})$ where

$$v(\mathbf{P}) = E[E(u, \mathbf{P}(s)), P^*] + \inf\{E[\mathbf{P}(s)\{x \geq 1 - \epsilon\}, P^*]:\epsilon > 0\},$$

with $\mathbf{P}(s)\{x \geq 1 - \epsilon\}$ the probability assigned by the simple measure $\mathbf{P}(s)$ to the subset $\{x:x \geq 1 - \epsilon, x \in X\}$ of X. Show that $B1$–$B5$ hold and that $B6$ and $B7$ do not hold.

15. (*Continuation.*) Let S, X, u, and P^* be as given in Exercise 14 with $P^*\{1, 3, 5, \ldots\} = P^*\{2, 4, 6, \ldots\} = 1/2$, and let $\mathbf{P} \prec \mathbf{Q} \Leftrightarrow v(\mathbf{P}) < v(\mathbf{Q})$ where

$$v(\mathbf{P}) = E[E(u, \mathbf{P}(s)), P^*] + \inf\{P^*\{E(u, \mathbf{P}(s)) \geq 1 - \epsilon\}:\epsilon > 0\}.$$

a. Prove that $(0 < \alpha < 1, \mathbf{P}, \mathbf{R} \in \mathcal{K}) \Rightarrow \inf\{P^*\{E(u, \alpha\mathbf{P}(s) + (1 - \alpha)\mathbf{R}(s)) \geq 1 - \epsilon\}:\epsilon > 0\} = \inf\{P^*(\{E(u, \mathbf{P}(s)) \geq 1 - \epsilon\} \cap \{E(u, \mathbf{R}(s)) \geq 1 - \epsilon\}):\epsilon > 0\}$. Note: $\{E(u, \mathbf{P}(s)) \geq 1 - \epsilon\} = \{s:E(u, \mathbf{P}(s)) \geq 1 - \epsilon\}$.
b. Show that $B1$, $B4$, $B5$, and $B7$ hold.
c. By specific example, show that $B2$ does not hold.
d. By specific example, show that $B3$ does not hold.
e. By specific example, show that $B6$ does not hold.

Note: In the remaining exercises, F is the set of all functions on S to X, $(f,g) \in F^2$ is interpreted as an even-chance alternative, x^* is the act in F that assigns $x \in X$ to every $s \in S$, and $A \subseteq S$ is null $\Leftrightarrow$ $(f,g) \sim (f',g')$ whenever $(f(s), g(s)) = (f'(s), g'(s))$ for all $s \in A^c$. Let $D1$ through $D7$ be the following axioms:

D1. $\prec$ *on* $F \times F = F^2$ *is a weak order.*
D2. $[(f,g') \preccurlyeq (f',g^*), (f',g) \preccurlyeq (f^*,g')] \Rightarrow (g,f) \preccurlyeq (g^*,f^*)$.
D3. $(X, \mathcal{T})$ *is a connected and separable topological space.*
D4. $\{(f,g):(f,g) \in F^2, (f,g) \prec (f',g')\} \in \mathcal{T}^{2n}$ *and* $\{(f,g):(f,g) \in F^2, (f',g') \prec (f,g)\} \in \mathcal{T}^{2n}$ *for each* $(f',g') \in F^2$.
D5. $(x^*, x^*) \prec (y^*, y^*)$ *for some* $x, y \in X$.
D6. (A *is not null;* $f = x, g = y, f' = z, g' = w$ *on* A; $\{f(s), g(s)\} = \{f'(s), g'(s)\}$ *for each* $s \in A^c$) $\Rightarrow [(x^*, y^*) \prec (z^*, w^*) \Leftrightarrow (f,g) \prec (f',g')]$.
D7. $[(f(s)^*, g(s)^*) \prec (f', g')$ *for all* $s \in S] \Rightarrow (f, g) \preccurlyeq (f',g')$; $[(f', g') \prec (f(s)^*, g(s)^*)$ *for all* $s \in S] \Rightarrow (f',g') \preccurlyeq (f,g)$.

16. Prove that if S is finite and $D1$–$D5$ hold then there is a real-valued function u on X and a probability measure P^* on S such that

$$(f,g) \prec (f',g') \Leftrightarrow E[u(f(s)), P^*] + E[u(g(s)), P^*] < E[u(f'(s)), P^*] + E[u(g'(s)), P^*]$$

for all (f,g), $(f',g') \in F^2$, with P^* unique and u unique up to a positive linear transformation when this holds.

17. Let S be of any size and assume that there is a real-valued function v on F that satisfies

$$(f,g) \prec (f',g') \Leftrightarrow v(f) + v(g) < v(f') + v(g'), \quad \text{for all} \quad (f,g), (f',g') \in F^2,$$

and is unique up to a positive linear transformation when it satisfies this. Assume also that the restriction of v to $\{x^*: x \in X\}$ is unique up to a positive linear transformation when it satisfies

$$(x^*, y^*) \prec (z^*, w^*) \Leftrightarrow v(x^*) + v(y^*) < v(z^*) + v(w^*), \qquad \text{for all } x, y, z, w \in X,$$

and that $D5$ and $D6$ hold.

Let $F_0 = \{f: f \in F \text{ and } \{f(s): s \in S\} \text{ is finite}\}$. Prove that, with $u(x) = v(x^*)$, there is a unique probability measure P^* on S such that

$$v(f) = E[u(f(s)), P^*] \tag{13.26}$$

for all $f \in F_0$, with A null $\Leftrightarrow P^*(A) = 0$. (Compare with $S1$.)

18. (*Continuation.*) Along with the assumptions of the preceding exercise assume that $D7$ holds and that for any $x, y \in X$ there is a $z \in X$ such that

$$(x^*, y^*) \sim (z^*, z^*). \tag{13.27}$$

Call $f \in F$ *bounded* $\Leftrightarrow P^*\{a \le u(f(s)) \le b\} = 1$ for some numbers a and b. Use the following steps to prove that (13.26) holds when f is bounded.

a. Show that $P^*\{a \le u(f(s)) \le b\} = 1 \Rightarrow a \le v(f) \le b$. [Let $A = \{s: a \le u(f(s)) \le b\}$, let $d = \sup\{u(f(s)): s \in A\}$, suppose $d < v(f)$ and use $D7$ to obtain a contradiction.]

b. With f bounded on A and $P^*(A) = 1$, for convenience assume $\inf\{u(f(s)): s \in A\} = 0$ and $\sup\{u(f(s)): s \in A\} = 1$, and assume (with no loss in generality since A^c is null) that $0 \le u(f(s)) \le 1$ on A^c. Given a positive integer n let $A_1 = \{s: 0 \le u(f(s)) \le 1/n\}$, $A_i = \{s: (i-1)/n < u(f(s)) \le i/n\}$ for $i = 2, \ldots, n$. By (13.27) there are $x_i \in X$ for which $(i-1)/n \le u(x_i) \le i/n$ for $i = 1, \ldots, n$. Define $f_i, g_i \in F$ by

$$f_i = f \text{ on } A_i; \qquad f_i = x_i \text{ on } A_i^c \qquad (i = 1, \ldots, n)$$
$$g_i = x_{i+1} \text{ on } \textstyle\bigcup_{j=1}^{i} A_j; \qquad g_i = x_i \text{ on } \bigcup_{j=i+1}^{n} A_j \qquad (i = 1, \ldots, n-1).$$

Use the fact that $\{f'(s), g'(s)\} = \{f''(s), g''(s)\}$ for all $s \in S$ implies $(f', g') \sim (f'', g'')$ along with the first $\Leftrightarrow$ expression in Exercise 17 to prove that

$$v(f) + \sum_{i=1}^{n-1} v(g_i) = \sum_{i=1}^{n} v(f_i). \tag{13.28}$$

c. Under the conditions in (*b*), it follows from (13.27) that for any $\epsilon > 0$ and any $i \in \{1, \ldots, n\}$ there are $x \in X$ with $|u(x) - i/n| < \epsilon$. Use this, (13.28), (13.26) for F_0, and the bounds on the $v(f_i)$ implied by step (*a*) to show that $\sum_i P^*(A_i)(i-1)/n - 1/n \le v(f) \le \sum_i P^*(A_i)i/n + 1/n$. Then argue that $v(f) = E[u(f(s)), P^*]$.

19. (*Continuation.*) Under the assumptions of the preceding exercise prove (*a*) u on X is bounded if there is a denumerable partition of S that has $P^*(A) > 0$ for every A in the partition, (*b*) every $f \in F$ is bounded.

Chapter 14

SAVAGE'S EXPECTED-UTILITY THEORY

The most brilliant axiomatic theory of utility ever developed is, in my opinion, the expected-utility theory of Savage (1954). It is an eminently suitable theory with which to conclude this book.

As has been true of significant developments throughout the history of mathematics, Savage's theory was not developed in a vacuum. He acknowledges and draws on the prior ideas of Ramsey (1931), de Finetti (1937), and von Neumann and Morgenstern (1947). His general approach is not unlike that presented by Ramsey in outline form. Unlike Ramsey, who proposed to first derive utility on the basis of an "ethically neutral proposition" or even-chance event and then to derive probabilities on the basis of utilities, Savage reverses this procedure. In his axiomatization of

$$f < g \Leftrightarrow E[u(f(s)), P^*] < E[u(g(s)), P^*], \qquad \text{for all } f, g \in F,$$

which is based solely on the binary relation $<$ on F, Savage first obtains the probability measure P^* on the set of subsets of S. This development owes much to de Finetti's work in probability theory. Using P^* Savage then obtains a structure much like that used by von Neumann and Morgenstern in their utility theory (Theorem 8.2), and proceeds to specify u on X. One final axiom then leads to the above representation on all of F. Savage's theorem is given in the next section which contains also an outline of later sections.

14.1 SAVAGE'S EXPECTED-UTILITY THEOREM

The main purpose of this chapter is to explore Theorem 14.1, which may appropriately be referred to as Savage's expected-utility theorem. This section presents the theorem and discusses its conditions and conclusions. Some preliminary definitions are required.

S is the set of states, X is the set of consequences, and F is the set of all functions on S into X. $A, B \subseteq S$; $x, y \in X$; $f, g \in F$. $\prec$ on F is the basic binary relation with $\sim$ and $\preccurlyeq$ defined in the usual way: $f \sim g \Leftrightarrow$ (not $f \prec g$, not $g \prec f$), and $f \preccurlyeq g \Leftrightarrow (f \prec g$ or $f \sim g)$.

$f = g$ on $A \Leftrightarrow f(s) = g(s)$ for all $s \in A$. $f = x$ on $A \Leftrightarrow f(s) = x$ for all $s \in A$. Partitions of S and complements are defined as in Section 13.1. A is *null* $\Leftrightarrow f \sim g$ whenever $f = g$ on A^c.

$x \prec y \Leftrightarrow f \prec g$ when $f = x$ and $g = y$ on S. $x \prec f \Leftrightarrow g \prec f$ when $g = x$ on S. Similar definitions hold for $x \sim y$, $f \sim y$, $x \preccurlyeq f$, and so forth.

Conditional preference is defined as follows: $f \prec g$ *given* $A \Leftrightarrow f' \prec g'$ whenever $f = f'$ and $g = g'$ on A, and $f' = g'$ on A^c. $\sim$ *given* A and $\preccurlyeq$ *given* A are defined in the usual way. $x \prec g$ *given* A means that $f \prec g$ given A whenever $f = x$ on A.

THEOREM 14.1. *Suppose that the following seven conditions hold for all* $f, g, f', g' \in F$, $A, B \subseteq S$, *and* $x, y, x', y' \in X$:

P1. $\prec$ *on F is a weak order;*

P2. $(f = f'$ *and* $g = g'$ *on* A, $f = g$ *and* $f' = g'$ *on* $A^c) \Rightarrow (f \prec g \Leftrightarrow f' \prec g')$;

P3. $(A$ *is not null*, $f = x$ *and* $g = y$ *on* $A) \Rightarrow (f \prec g$ *given* $A \Leftrightarrow x \prec y)$;

P4. $[(x \prec y$, $f = y$ *on* A, $f = x$ *on* A^c, $g = y$ *on* B, $g = x$ *on* $B^c)$ *and* $(x' \prec y'$, $f' = y'$ *on* A, $f' = x'$ *on* A^c, $g' = y'$ *on* B, $g' = x'$ *on* $B^c)] \Rightarrow (f \prec g \Leftrightarrow f' \prec g')$;

P5. $x \prec y$ *for some* $x, y \in X$;

P6. $(f \prec g, x \in X) \Rightarrow$ *there is a finite partition of S such that, if A is any event in the partition, then* $(f' = x$ *on* A, $f' = f$ *on* $A^c) \Rightarrow f' \prec g$, *and* $(g' = x$ *on* A, $g' = g$ *on* $A^c) \Rightarrow f \prec g'$;

P7. $(f \prec g(s)$ *given* A, *for all* $s \in A) \Rightarrow f \preccurlyeq g$ *given* A. $(g(s) \prec f$ *given* A, *for all* $s \in A) \Rightarrow g \preccurlyeq f$ *given* A.

Then, with $\prec^*$ *defined on the set of all subsets of S by*

$$A \prec^* B \Leftrightarrow f \prec g \quad \textit{whenever } (x \prec y, f = y \textit{ on } A, f = x \textit{ on } A^c, g = y \textit{ on } B, g = x \textit{ on } B^c), \quad (14.1)$$

there is a unique probability measure P^* *on the set of all subsets of S that satisfies*

$$A \prec^* B \Leftrightarrow P^*(A) < P^*(B), \qquad \textit{for all } A, B \subseteq S, \qquad (14.2)$$

and P^* *has the property that*

$$(B \subseteq S, 0 \leq \rho \leq 1) \Rightarrow P^*(C) = \rho P^*(B) \qquad \textit{for some } C \subseteq B; \quad (14.3)$$

and, with P^* *as given, there is a real-valued function u on X for which*

$$f \prec g \Leftrightarrow E[u(f(s)), P^*] < E[u(g(s)), P^*], \qquad \textit{for all } f, g \in F, \quad (14.4)$$

and when u satisfies this it is bounded and unique up to a positive linear transformation.

The final condition, $P7$, is similar to $A4a$ of Section 10.4 and $B6$ of Section 13.3. It is an obvious dominance (or sure-thing, or independence) condition and it is *not* required in the derivation of P^* that satisfies (14.2), just as $B6$ was not required in the derivation of P^* for Theorem 13.3. The form of $P7$ given in the theorem is slightly weaker than (does not assume as much as) Savage's original form which has $\preccurlyeq$ where $<$ appears in $P7$, but the two are equivalent in the presence of the other conditions, $P1$–$P6$.

$P2$ and $P3$ explicate Savage's "sure-thing principle." $P2$ says that preferences between acts should not depend on those states that have identical consequences for the two acts. It is closely related to the independence condition of Chapter 8 and is found reasonable by many persons provided that the state that obtains does not depend on the act that is actually implemented. Together, $P1$ and $P2$ imply that $<$ given A is a weak order on F for every $A \subseteq S$.

$P3$, as a companion to $P2$, says that if $f = x$ and $g = y$ on A and if A is not null, then $f < g$ given $A \Leftrightarrow f' < g'$ when $f' = x$ and $g' = y$ on S. This sets up a reasonable correspondence between preferences on consequences (constant acts) and conditional preferences on events that the decision maker regards as possible.

$<^*$ as defined in (14.1) is a *qualitative probability relation* on the set of events. We read $A <^* B$ as "A is less probable than B." As noted in (14.1), "is less probable than" is defined in terms of "is less preferred than." The principle objective of $P4$ in this connection is to ensure that $<^*$ on the events is a weak order. Suppose you prefer y to x and can either take your chances on getting y if A obtains or on getting y if B obtains. In either case if the event you take your chances on does not obtain you will receive the less preferred x. If you select B then it seems reasonable to suppose that you regard B as more probable than A. $P4$ says that if you prefer to take your chances on getting y if B obtains then, with two other consequences y' and x' with y' preferred to x', you would (or ought to) rather take your chances on getting y' if B obtains than on getting y' if A obtains. As in the case of $P2$, this seems reasonable as long as the state that obtains does not depend on the consequences assigned to the states by any particular act.

$P5$ says that indifference does not hold between every pair of constant acts. It is needed to ensure the uniqueness of P^*. If $P5$ were false then $<^*$ would be reflexive. For further remarks see Section 14.3.

The effect of $P6$, which is a rather strong assumption, can best be seen from (14.3) which in the presence of (14.2) follows from $P1$–$P6$. Among other things, (14.3) says that S must be uncountable, that $P^*(s) = 0$ for every $s \in S$, and that for any positive integer n there is an n-event partition of S

having $P^*(A) = 1/m$ for each A in the partition. We shall refer to such partitions as *uniform partitions*.

In guaranteeing $A <^* B \Leftrightarrow P^*(A) < P^*(B)$, $P6$ has an unmistakable Archimedean quality. In effect it says that no consequence is "infinitely desirable" (which would negate $f' < g$ if x were so desirable) and that no consequence is "infinitely undesirable" (which would negate $f < g'$ if x were so undesirable). If S is allowed to be infinite, something like $P6$ is required to ensure the existence of real-valued order ($<^*$) preserving probabilities. As Savage points out, weaker versions of $P6$ are sufficient for (14.2), but may not yield (14.3) as well. The usefulness of (14.3) will become apparent when we see how it is used as a point of departure in defining gambles on X that lead to the definition of the utility function u on X.

$P1$ through $P6$ are sufficient to obtain (14.4) for all acts in F that assign no more than a finite number of consequences to all the states in some event A for which $P^*(A) = 1$. $P7$ is then used (as was $B6$ in the preceding chapter) to verify that (14.4) holds for *all* acts, and it ensures that u on X is bounded. When he wrote *The Foundations of Statistics*, Savage had the impression that $P1$–$P7$ do not imply that u is bounded. Some years later, when we were working on the theory that appears in Chapter 10 of Part II, we discovered that this impression was false. Because of the false impression, Savage did not state (14.4) for all acts but, in light of the boundedness of u, he did in fact prove (14.4) as it is presented here. In other words, he proved (14.4) for all bounded acts. Since u is bounded, all acts are bounded. The proof of boundedness given later is essentially his.

In proving Theorem 14.1 I shall follow the pattern used by Savage. Here is a sectional outline.

Section 14.2 shows that (14.2) and (14.3) follow from five conditions ($F1$–$F5$) for $<^*$ on the set of events.

Section 14.3 under definition (14.1) shows that $P1$–$P6 \Rightarrow F1$–$F5$.

Section 14.4 establishes, from $P1$–$P6$, the three preference axioms of Theorem 8.2, which shows that (14.4) holds for acts confined with probability one to a finite subset of consequences.

Section 14.5 proves that u on X is bounded. This uses $P7$.

Section 14.6 uses $P7$ to verify (14.4) for all acts.

The proofs that follow are essentially Savage's. I have added some details to them in places where I felt that this would aid some readers.

14.2 AXIOMS FOR PROBABILITY

In this section we shall prove the following theorem.

THEOREM 14.2. *Suppose that $<^*$ on the set of all subsets of S satisfies the*

following conditions for all A, B, $C \subseteq S$:

$F1$. *not* $A <^* \varnothing$,
$F2$. $\varnothing <^* S$,
$F3$. $<^*$ *is a weak order*,
$F4$. $A \cap C = B \cap C = \varnothing \Rightarrow (A <^* B \Leftrightarrow A \cup C <^* B \cup C)$,
$F5$. $A <^* B \Rightarrow$ *there is a finite partition* $\{C_1, \ldots, C_m\}$ *of* S *for which* $A \cup C_i <^* B$ *for* $i = 1, \ldots, m$.

Then there is one and only one probability measure P^* on the set of all subsets of S that satisfies (14.2), *and* (14.3) *holds for this measure.*

$F1$–$F4$, which define $<^*$ as a *qualitative probability*, are necessary for (14.2), but collectively they are not sufficient. $F1$–$F5$ as noted are sufficient but $F5$ is not necessary for (14.2) although it does follow from (14.2) and (14.3).

As usual we define $A \sim^* B \Leftrightarrow$ (not $A <^* B$, not $B <^* A$), and $A \leqslant^* B \Leftrightarrow$ ($A <^* B$ or $A \sim^* B$). Throughout this section and the rest of this chapter, I shall use A/B ("A but not B") to denote the complement of B relative to A:

$$A/B = A \cap B^c. \tag{14.5}$$

In approaching Theorem 14.2 we begin with a series of consequences of $F1$–$F5$. $C1$ through $C4$ presuppose only $F1$–$F4$. The rest presuppose all of $F1$ through $F5$.

$C1$. $B \subseteq C \Rightarrow \varnothing \leqslant^* B \leqslant^* C \leqslant^* S$.
$C2(\sim^*)$. $(A \sim^* B, B \cap C = \varnothing) \Rightarrow A \cup C \leqslant^* B \cup C$.
$C2(<^*)$. $(A <^* B, B \cap C = \varnothing) \Rightarrow A \cup C <^* B \cup C$.
$C3(\sim^*)$. $(A \sim^* B, C \sim^* D, B \cap D = \varnothing) \Rightarrow A \cup C \leqslant^* B \cup D$.
$C3(<^*)$. $(A \leqslant^* B, C <^* D, B \cap D = \varnothing) \Rightarrow A \cup C <^* B \cup D$.
$C4$. $(A \sim^* B, C \sim^* D, A \cap C = B \cap D = \varnothing) \Rightarrow A \cup C \sim^* B \cup D$.
$C5$. $\varnothing <^* A \Rightarrow A$ *can be partitioned into two events B and C for which* $(\varnothing <^* B, \varnothing <^* C)$.
$C6$. (A, B, *and C are pairwise disjoint*, $A \leqslant^* B$, $B <^* A \cup C) \Rightarrow$ *there is a* $D \subseteq C$ *for which* $\varnothing <^* D$ *and* $B \cup D <^* A \cup (C/D)$.
$C7$. $(\varnothing <^* A, \varnothing <^* B, A \cap B = \varnothing) \Rightarrow B$ *can be partitioned into C and D for which* $C \leqslant^* D \leqslant^* A \cup C$.
$C8$. $\varnothing <^* A \Rightarrow A$ *can be partitioned into B and C with* $B \sim^* C$.
$C9$. $\varnothing <^* A \Rightarrow$ *for any positive integer n there is a 2^n part partition of A such that $\sim^*$ holds between each two events in the partition.*

Proofs of $C1$ through $C9$

$C1$. The proof is easy and is left to the reader.

$C2(\sim^*)$. Assume $(A \sim^* B, B \cap C = \varnothing)$. Since $A = (A/C) \cup (A \cap C)$ and $A \cap (C/A) = \varnothing$, $F4 \Rightarrow (A/C) \cup (A \cap C) \cup (C/A) \sim^* B \cup (C/A)$, or $A \cup C \sim^* B \cup (C/A)$. By $C1$, $B \cup (C/A) \leqslant^* B \cup C$. Hence, by $F3$, $A \cup C \leqslant^* B \cup C$.

$C2(<^*)$. Replace $\sim^*$ by $<^*$ in preceding proof.

$C3(\sim^*)$. Assume $(A \sim^* B, C \sim^* D, B \cap D = \varnothing)$. Since $(C/B) \cap B = \varnothing$, $C2(\sim^*) \Rightarrow A \cup (C/B) \leqslant^* B \cup (C/B) = B \cup C$. Also, since $(B/C) \cap D = \varnothing$, $C2(\sim^*)$ and $C \sim^* D$ imply $B \cup C = C \cup (B/C) \leqslant^* D \cup (B/C)$. By $F3$, $A \cup (C/B) \leqslant^* D \cup (B/C)$. This, $C2$, and $(B \cap C) \cap (D \cup (B/C)) = \varnothing$ then imply that $A \cup (C/B) \cup (B \cap C) \leqslant^* D \cup (B/C) \cup (B \cap C)$, or $A \cup C \leqslant^* D \cup B$.

$C3(<^*)$. Replace $C \sim^* D$ by $C <^* D$ in preceding proof. Use $C2(<^*)$.

$C4$. Assume $(A \sim^* B, C \sim^* D, A \cap C = B \cap D = \varnothing)$. By $C3(\sim^*)$, $A \cup C \leqslant^* B \cup D$ and $B \cup D \leqslant^* A \cup C$. Hence $A \cup C \sim^* B \cup D$.

$C5$. Assume $\varnothing <^* A$. $F5 \Rightarrow$ there is a partition $\{D_1, \ldots, D_m\}$ of S for which $D_i <^* A$ for each i. $C1 \Rightarrow D_i \cap A \leqslant^* (D_i \cap A) \cup (D_i/A) = D_i$. Hence $D_i \cap A <^* A$ for all i. If $D_i \cap A \sim^* \varnothing$ for each i then, by $C4$, $\bigcup_i (D_i \cap A) \sim^* \varnothing$, or $A \sim^* \varnothing$, a contradiction. If $\varnothing <^* D_i \cap A$ for only one i, say $i = 1$, then $A \sim^* D_1 \cap A$ which contradicts $D_1 \cap A <^* A$. Hence $\varnothing <^* D_i \cap A$ for at least two i.

$C6$. Assume $(A \cap B = A \cap C = B \cap C = \varnothing, A \leqslant^* B, B <^* A \cup C)$. $(F3, F4) \Rightarrow \varnothing <^* C$. Since $B <^* A \cup C$ and $\varnothing <^* C$, it follows from $F5$ that there is a $D_1 \subseteq C$ for which $\varnothing <^* D_1$ and $B \cup D_1 <^* A \cup C$. By $C5$ and $F3$, D_1 can be partitioned into D and D' with $\varnothing <^* D \leqslant^* D'$, so that $B \cup D \cup D' <^* A \cup (C/D) \cup D$. $F4 \Rightarrow B \cup D' <^* A \cup (C/D)$. $(F4, D \leqslant^* D') \Rightarrow B \cup D \leqslant^* B \cup D'$. Hence $B \cup D <^* A \cup (C/D)$.

$C7$. Assume $(\varnothing <^* A, \varnothing <^* B, A \cap B = \varnothing)$. If $B \leqslant^* A$ the conclusion follows easily from $C5$. Assume that $A <^* B$. $F5 \Rightarrow$ there is a partition $\{G_1, \ldots, G_n\}$ of B such that $G_i <^* A$ for each i. For definiteness assume that $G_1 \leqslant^* \cdots \leqslant^* G_n$. Let m be such that $\bigcup_1^m G_i \leqslant^* \bigcup_{m+1}^n G_i \leqslant^* \bigcup_1^{m+1} G_i$. Let $C = \bigcup_1^m G_i$ and $D = \bigcup_{m+1}^n G_i$. Then $C \leqslant^* D \leqslant^* C \cup G_{m+1}$ which, since $G_{m+1} <^* A$, implies by $F4$ and $F3$ that $D <^* C \cup A$.

$C8$. Assume $\varnothing <^* A$. It follows from $C5$ that A can be partitioned into B_1, C_1, D_1 such that $B_1 \leqslant^* C_1 \cup D_1$ and $C_1 \leqslant^* B_1 \cup D_1$. If one of these two $\leqslant^*$ is $\sim^*$, the conclusion of $C8$ holds. Henceforth assume that $B_1 <^* C_1 \cup D_1$ and $C_1 <^* B_1 \cup D_1$. Then $\varnothing <^* D_1$. For definiteness take $B_1 \leqslant^* C_1$. Then $C6 \Rightarrow$ there is a $C^2 \subseteq D_1$ such that $\varnothing <^* C^2$ and $C_1 \cup C^2 \leqslant^* B_1 \cup (D_1/C^2)$. Hence $\varnothing <^* D_1/C^2$ and, by $C7$, D_1/C^2 can be partitioned into B^2 and D_2 such that $B^2 \leqslant^* D_2 \leqslant^* C^2 \cup B^2$. Since $B_1 \leqslant^* C_1$, $B_1 \cup B^2 \leqslant^* C_1 \cup D_2 <^* C_1 \cup D_2 \cup C^2$, all by $F4$. Let

$$B_2 = B_1 \cup B^2 \quad \text{and} \quad C_2 = C_1 \cup C^2.$$

We then obtain a partition $\{B_2, C_2, D_2\}$ of A for which

1. $B_2 <^* C_2 \cup D_2$ and $C_2 <^* B_2 \cup D_2$,
2. $B_1 \subseteq B_2$, $C_1 \subseteq C_2$, $D_2 \subseteq D_1$,
3. $D_2 \preccurlyeq^* D_1/D_2$.

By repeating this process it follows that there is a sequence $\ldots, \{B_n, C_n, D_n\}, \ldots$ of three part partitions of A such that, for each $n \geq 1$,

1. $B_n <^* C_n \cup D_n$ and $C_n <^* B_n \cup D_n$,
2. $B_n \subseteq B_{n+1}$, $C_n \subseteq C_{n+1}$, $D_{n+1} \subseteq D_n$,
3. $D_{n+1} \preccurlyeq^* D_n/D_{n+1}$,

so that $\varnothing <^* D_n$ for all n, and D_n contains two disjoint events (D_{n+1} and D_n/D_{n+1}) each of which is as probable as D_{n+1}. Hence, using (3) and $C3(<^*)$, $(E_1 <^* D_{n+1}, E_2 <^* D_{n+1}) \Rightarrow E_1 \cup E_2 <^* D_n$.

Now for any G with $\varnothing <^* G$, $D_n <^* G$ for sufficiently large n. For example, if $G \preccurlyeq^* D_n$ then, with $\{E_1, \ldots, E_m\}$ for $\varnothing <^* G$ as in $F5$ with $E_i <^* G$ for all i, $E_i <^* D_n$ for all i so that $E_1 \cup E_2 <^* D_{n-1}$, $E_3 \cup E_4 <^* D_{n-1}, \ldots$ and then $\bigcup_1^4 E_i <^* D_{n-2}$, $\bigcup_5^8 E_i <^* D_{n-2}, \ldots$ and so forth, so that with n sufficiently large $\bigcup_{i=1}^m E_i <^* D_1$, or $S <^* D_1$, which is false. In addition, $\varnothing \sim^* \bigcap_{n=1}^\infty D_n$, for if $\varnothing <^* \bigcap_{n=1}^\infty D_n$ then $D_m <^* \bigcap D_n$ for sufficiently large m, and this is false since $\bigcap D_n \subseteq D_m$.

Let

$$B = \bigcup_{n=1}^{\infty} B_n \quad \text{and} \quad C = \left(\bigcup_{n=1}^{\infty} C_n\right) \cup \left(\bigcap_{n=1}^{\infty} D_n\right).$$

$\{B, C\}$ is a partition of A since $(\bigcup B_n) \cap (\bigcup C_n) = (\bigcup B_n) \cap (\bigcap D_n) = (\bigcup C_n) \cap (\bigcap D_n) = \varnothing$. To verify $B \sim^* C$ note first that $C \sim^* \bigcup C_n$ since $\bigcap D_n \sim^* \varnothing$. Suppose that $B <^* C$. Then $B <^* \bigcup C_n$ and, by $C6$, there is a $G \subseteq \bigcup C_n$ for which $\varnothing <^* G$ and

$$B \cup G <^* (\bigcup C_n)/G.$$

Since $B \cap G = \varnothing$ and $B_n \preccurlyeq^* B$ (since $B_n \subseteq B$), $F4$ implies

$$B_n \cup G \preccurlyeq^* B \cup G.$$

For large n $D_n <^* G$ so that, again by $F4$,

$$B_n \cup D_n <^* B_n \cup G.$$

Since $D_n \cap (\bigcup C_n) <^* G$ for large n and $\bigcup C_n = (\bigcup C_n/G) \cup G = (\bigcup C_n/D_n) \cup ((\bigcup C_n) \cap D_n)$, it follows by $C3(<^*)$ that for large n

$$\bigcup C_n/G \preccurlyeq^* (\bigcup C_n)/D_n.$$

Finally, since $(\bigcup C_n)/D_n \subseteq C_n$, $C1 \Rightarrow (\bigcup C_n)/D_n \preccurlyeq^* C_n$. This and the four preceding displayed expressions yield $B_n \cup D_n <^* C_n$ by transitivity (for large n) which contradicts $C_n <^* B_n \cup D_n$ in (1). Therefore, not $B <^* C$. By a similar proof, not $\bigcup C_n <^* (\bigcup B_n) \cup (\bigcap D_n)$ so that not $C <^* B$.

C9. This follows from $C3(<^*)$ and $C8$. ◆

We now complete the proof of Theorem 14.2.

Proof of (14.2)

Let $F1$–$F5$ hold. We shall call a partition $\{A_1, \ldots, A_m\}$ of A a u.p. (*uniform partition*) when $\varnothing <^* A$ and $A_1 \sim^* A_2 \sim^* \cdots \sim^* A_m$, and let

$$C(r, 2^n) = \{A : A \text{ is the union of } r \text{ events in some } 2^n \text{ part u.p. of } S\}.$$

We shall establish (14.2) through a series of steps, each of which proves a key assertion.

1. $[A, B \in C(r, 2^n)] \Rightarrow A \sim^* B$. First, if $A, B \in C(1, 2^n)$, and if $A <^* B$, it follows easily from $C3(<^*)$ that $S <^* S$. Hence, if $A, B \in C(1, 2^n)$, then $A \sim^* B$. Therefore, if $A, B \in C(r, 2^n)$, $A \sim^* B$ follows from $C4$.

2. $[A \in C(r, 2^n),\ B \in C(r2^m, 2^{n+m})] \Rightarrow A \sim^* B$. First, if $A \in C(1, 2^n)$ and $B \in C(2^m, 2^{n+m})$, then $A \sim^* B$, for otherwise, by step 1 and $C3(<^*)$ we get $S <^* S$. The desired conclusion follows from $C4$.

3. $[A \in C(r, 2^n), B \in C(t, 2^m)] \Rightarrow (A \preccurlyeq^* B \Leftrightarrow r/2^n \le t/2^m)$. If $r/2^n = t/2^m$ then $r2^m = t2^n$ and, with $D \in C(r2^m, 2^{n+m})$ it follows from step 2 that $A \sim^* D$ and $B \sim^* D$, so that $A \sim^* B$. If $r2^m < t2^n$ then, with $D_1 \in C(r2^m, 2^{n+m})$ and $D_2 \in C(t2^n, 2^{n+m})$ we get $A \sim^* D_1$ and $B \sim^* D_2$. But surely $D_1 <^* D_2$ when $r2^m < t2^n$. Therefore $A <^* B$.

4. For $A \subseteq S$ let $k(A, 2^n)$ be the *largest* integer r (possibly zero) such that $B \preccurlyeq^* A$ when $B \in C(r, 2^n)$, and define

$$P^*(A) = \sup\{k(A, 2^n)/2^n : n = 0, 1, 2, \ldots\}. \tag{14.6}$$

Clearly, $P^*(\varnothing) = 0$, $P^*(S) = 1$, and $P^*(A) \ge 0$ for all $A \subseteq S$. Moreover,

$$A \in C(r, 2^n) \Rightarrow P^*(A) = r/2^n. \tag{14.7}$$

If $A \in C(r, 2^n)$ then, by (14.6), $P^*(A) \ge r/2^n$. If, in fact $P^*(A) > r/2^n$ then for some $B \in C(t, 2^m)$ with $r/2^n < t/2^m$, $B \preccurlyeq^* A$. But this is impossible by step 3.

5. $A \preccurlyeq^* B \Rightarrow P^*(A) \le P^*(B)$. This is obvious from (14.6).

6. *P^* is finitely additive*. Let $A \cap B = \varnothing$. It follows that, for each n, there is a 2^n part u.p. of S for which A_n and B_n are unions of elements in this partition, with $A_n \cap B_n = \varnothing$, $A_n \in C(k(A, 2^n), 2^n)$, $B_n \in C(k(B, 2^n), 2^n)$, $A_n \preccurlyeq^* A$, $B_n \preccurlyeq^* B$. Hence $A_n \cup B_n \preccurlyeq^* A \cup B$ by $C3$, and $k(A, 2^n) + k(B, 2^n) \le k(A \cup B, 2^n)$. Since, for any $A \subseteq S$, it is easily seen that

$k(A, 2^n)/2^n$ does not decrease as n increases, it follows from Exercise 10.7 that

$$P^*(A) + P^*(B) \leq P^*(A \cup B).$$

If we now define $k^*(A, 2^n)$ as the *smallest* integer r such that $A \preccurlyeq^* B$ when $B \in C(r, 2^n)$, it readily follows from the fact that $\{r/2^n : r = 0, \ldots, 2^n; n = 0, 1, \ldots\}$ is dense in $[0, 1]$ that $\inf\{k^*(A, 2^n)/2^n : n = 0, 1, \ldots\} = \sup\{k(A, 2^n)/2^n : n = 0, 1, \ldots\}$. A proof symmetric to that just completed then implies that

$$P^*(A \cup B) \leq P^*(A) + P^*(B) \qquad \text{when } A \cap B = \varnothing$$

so that $P^*(A \cup B) = P^*(A) + P^*(B)$.

7. $\varnothing <^* A \Rightarrow 0 < P^*(A)$. Let $\varnothing <^* A$. By *F5* there is a partition $\{A_1, \ldots, A_n\}$ of S for which $A_i <^* A$ for each i. Then, by step 5, $P^*(A_i) \leq P^*(A)$. Finite additivity then requires that $P^*(A) > 0$.

8. $A <^* B \Rightarrow P^*(A) < P^*(B)$. Suppose $A <^* B$. Then, using *F5*, there is a $C \subseteq S$ for which $\varnothing <^* C$, $C \cap A = \varnothing$, and $C \cup A <^* B$. By finite additivity and step 5, $P^*(C) + P^*(A) \leq P^*(B)$. Since $P^*(C) > 0$ by step 7, $P^*(A) < P^*(B)$.

Steps 5 and 8 imply (14.2) and it is obvious that P^* as defined here is the only probability measure on S that satisfies (14.2). ◆

Proof that $(B \subseteq S,\ 0 \leq \rho \leq 1) \Rightarrow P^*(C) = \rho P^*(B)$ **for some** $C \subseteq B$

If $P^*(B) = 0$ the result is obvious. Assume then that $P^*(B) > 0$, and consider a sequence $\{A_1^1, A_2^1\}, \{A_1^2, \ldots, A_4^2\}, \ldots, \{A_1^n, \ldots, A_{2^n}^n\}, \ldots$ of 2^n part u.p.'s of B for which $\{A_{2i-1}^{n+1}, A_{2i}^{n+1}\}$ is a 2 part u.p. of A_i^n. For a given n let $m = \sup\{j : P^*(\bigcup_{i=1}^j A_i^n) < \rho P^*(B)\}$ so that

$$P^*\left(\bigcup_1^m A_i^n\right) + 2^{-n}P^*(B) \geq \rho P^*(B),$$

and let $k = \inf\{j : P^*(\bigcup_{i=j}^{2^n} A_i^n) < (1 - \rho)P^*(B)\}$ so that

$$P^*\left(\bigcup_k^{2^n} A_i^n\right) + 2^{-n}P^*(B) \geq (1 - \rho)P^*(B).$$

Let $C_n = \bigcup_{i=1}^m A_i^n$ and $D_n = \bigcup_{i=k}^{2^n} A_i^n$ so that $C_1 \subseteq C_2 \subseteq \cdots$, $D_1 \subseteq D_2 \subseteq \cdots$, $C_n \cap D_n = \varnothing$ for all n, and $P^*(C_n) \geq \rho P^*(B) - 2^{-n}P^*(B)$ and $P^*(D_n) \geq (1 - \rho)P^*(B) - 2^{-n}P^*(B)$ for all n. Since $C_n \subseteq \bigcup_n C_n$ and $D_n \subseteq \bigcup_n D_n$, $\rho P^*(B) \leq P^*(\bigcup C_n)$ and $(1 - \rho)P^*(B) \leq P^*(\bigcup_n D_n)$. Moreover, $(\bigcup C_n) \cap (\bigcup D_n) = \varnothing$. Hence, by finite additivity, *C*1, and (14.2),

$$P^*(\bigcup C_n) + P^*(\bigcup D_n) = P^*((\bigcup C_n) \cup (\bigcup D_n)) \leq P^*(B)$$

which requires $P^*(\bigcup C_n) = \rho P^*(B)$ and $P^*(\bigcup D_n) = (1 - \rho)P^*(B)$. ◆

14.3 PROBABILITIES FROM PREFERENCES

This section shows how $F1$–$F5$ of Theorem 14.2 follow from $P1$–$P6$ and (14.1), which is

$$A <^* B \Leftrightarrow [(x < y, f = y \text{ on } A, f = x \text{ on } A^c, g = y \text{ on } B, g = x \text{ on } B^c) \Rightarrow f < g]. \quad (14.1)$$

If $x \sim y$ for all $x, y \in X$ then $A <^* B$ for every $A, B \subseteq S$. $P5$ clears up this potential snag. [Savage, who uses a different definition than (14.1), gets $A \sim^* B$ for all $A, B \subseteq S$ when $P5$ is false. His definition is $A \leqslant^* B \Leftrightarrow [(x < y, \ldots) \Rightarrow f \leqslant g]$. The main difference here is stylistic.]

Since $P5$ says that $x < y$ for some $x, y \in X$, it then follows from (14.1) and $P1$ that $<^*$ is asymmetric: $A <^* B \Rightarrow$ not $B <^* A$. Suppose not $A <^* B$ and not $B <^* C$. With $x < y$ it follows from $P4$ and (14.1) that ($f = y$ on A, $f = x$ on A^c, $g = y$ on B, $g = x$ on B^c, not $f < g$) and that ($g = y$ on B, $g = x$ on B^c, $h = y$ on C, $h = x$ on C^c, not $g < h$), so that, using $P1$, ($f = y$ on A, $f = x$ on A^c, $h = y$ on C, $h = x$ on C^c, not $f < h$), so that not $A <^* C$. Hence $(P1, P4, P5) \Rightarrow F3$. $<^*$ on the set of all subsets of S is a weak order.

Letting $A = \varnothing$ and $B = S$ in (14.1), $\varnothing <^* S$ follows immediately from the definition of $<$ on X. $\varnothing <^* S$ is $F2$.

Suppose A is null and ($x < y, f = y$ on A, $f = x$ on A^c, $g = x$ on S). Then, since $f = g$ on A^c, $f \sim g$. Hence not $A <^* \varnothing$. If A is not null and ($x < y, f = x$ on S, $g = y$ on A, $g = x$ on A^c) then $f < g$ given A by $P3$, and since $f = g$ on A^c, $f < g$ by the definition of conditional preference. It then follows that $\varnothing <^* A$. This verifies $F1$ in the presence of $F3$.

$F4$ is implied by $P2$ and $P4$. Assume $A \cap C = B \cap C = \varnothing$. If $P5$ is false then $A <^* B$ and $A \cup C <^* B \cup C$ follow. Assume then that $x < y$. Let

$$\begin{array}{llll} f = y & \text{on } A, & f = x & \text{on } A^c \\ g = y & \text{on } B, & g = x & \text{on } B^c \\ f' = y & \text{on } A \cup C, & f' = x & \text{on } (A \cup C)^c \\ g' = y & \text{on } B \cup C, & g' = x & \text{on } (B \cup C)^c. \end{array}$$

Since $f = f'$ and $g = g'$ on C^c, and $f = g$ and $f' = g'$ on C, $P2$ says that $f < g \Leftrightarrow f' < g'$. If $A <^* B$ then $f < g$ by (14.1), then $f' < g'$, then $A \cup C <^* B \cup C$ by (14.1) and $P4$. By the reverse procedure $A \cup C <^* B \cup C \Rightarrow A <^* B$.

To verify $F5$ suppose $A <^* B$. Take $x < y$ by $P5$. With f, g as in (14.1), $f < g$. By $P6$ there is a partition $\{C_1, \ldots, C_m\}$ of S such that $f_i < g$ when $f_i = y$ on C_i and $f_i = f$ on C_i^c. Since $f_i = y$ on $A \cup C_i$ and $f_i = x$ on $(A \cup C_i)^c$ and $f_i < g$, (14.1) and $P4$ imply $A \cup C_i <^* B$.

Thus, $F1$–$F5$ follow from $P1$–$P6$ under (14.1). Therefore, by Theorem 14.2, $P1$–$P6$ imply the existence of P^* as specified in (14.2) and (14.3).

14.4 UTILITY FOR SIMPLE ACTS

P^* as specified in (14.2) and (14.3) induces a probability measure P_f on (the set of subsets of) X for each $f \in F$ as follows:

$$P_f(Y) = P^*\{f(s) \in Y\} \quad \text{for each } Y \subseteq X, \tag{14.8}$$

where, as usual, $P^*\{f(s) \in Y\}$ means $P^*(\{s : f(s) \in Y\})$. Let $\mathcal{P}_s$ be the set of all simple probability measures on X and let $\mathcal{P} = \{P_f : f \in F\}$. With F the set of all functions on S to X it follows from (14.3) that $\mathcal{P}_s \subseteq \mathcal{P}$.

Later in this section we shall prove that the three conditions of Theorem 8.2 follow from $P1$–$P6$. Before doing that we note that for any $P \in \mathcal{P}$ there may be many different acts in F that have this P as their measure on X induced by P^*. Clearly then, if (14.4) is to hold it is absolutely essential to have $f \sim g$ when $P_f = P_g$.

THEOREM 14.3. $(P1\text{–}P6; P_f = P_g; P_f, P_g \in \mathcal{P}_s) \Rightarrow f \sim g$.

Preparatory to proving this we shall prove two lemmas, the first of which will be used extensively in later developments.

LEMMA 14.1. *$(P1, P2, \{A_1, \ldots, A_n\}$ is a partition of A, $f \lesssim g$ given A_i for each $i) \Rightarrow f \lesssim g$ given A. $(P1, P2, \{A_1, \ldots, A_n\}$ is a partition of A, $f \lesssim g$ given A_i for each i, $f < g$ given A_i for some $i) \Rightarrow f < g$ given A.*

LEMMA 14.2. *$(P1$–$P4$, $A \cap B = \varnothing$, $A \sim^* B$, $f = x$ and $g = y$ on A, $f = y$ and $g = x$ on $B) \Rightarrow f \sim g$ given $A \cup B$.*

Proof of Lemma 14.1. Let the hypotheses of the first part hold. Let $f' = f$ and $g' = g$ on A, and $f' = g'$ on A^c. By $(P1, P2)$, $f \lesssim g$ given $A \Leftrightarrow f' \lesssim g'$. For $i = 1, \ldots, n-1$ let

$$\begin{aligned} f_i &= g' = g \quad \text{on} \quad \bigcup_{j=1}^{i} A_j \\ f_i &= f' = f \quad \text{on} \quad \bigcup_{j=i+1}^{n} A_j \\ f_i &= f' = g' \quad \text{on} \quad A^c. \end{aligned}$$

Since $f \lesssim g$ given A_i for each i, $(P1, P2) \Rightarrow f' \lesssim f_1, f_1 \lesssim f_2, \ldots, f_{n-1} \lesssim g'$ and hence $f' \lesssim g'$. If $f < g$ given A_i for some i also then one $\lesssim$ in the sequence is $<$ and hence $f' < g'$, or $f < g$ given A. ◆

Proof of Lemma 14.2. Let the hypotheses of the lemma hold. Let

$$f' = y \quad \text{on} \quad B, \qquad f' = x \quad \text{on} \quad B^c$$
$$g' = y \quad \text{on} \quad A, \qquad g' = x \quad \text{on} \quad A^c.$$

If $x < y$ then $f' \sim g'$ by $A \sim^* B$, (14.1), and $P4$. Since $f' = g' = x$ on $(A \cup B)^c$, $P1 \Rightarrow f' \sim g'$ given $(A \cup B)^c$. Then $f' \sim g'$ given $A \cup B$ for otherwise, by Lemma 14.1, either $f' < g'$ or $g' < f'$. Since $f = f'$ and $g = g'$ on $A \cup B$, $(P1, P2) \Rightarrow f \sim g$ given $A \cup B$. If $y < x$ the conclusion is the same. Finally, suppose $x \sim y$. If $A(B)$ is null then $f \sim g$ given $A(B)$ follows from the definitions of conditional preference and null events. If $A(B)$ is not null then $f \sim g$ given $A(B)$ follows directly from $P3$. Hence, by Lemma 14.1, $f \sim g$ given $A \cup B$. ◆

Proof of Theorem 14.3. Let $P1$–$P6$ hold. We are to prove that if the x_i are all different and if

$$f = x_i \quad \text{on} \quad A_i, \; g = x_i \quad \text{on} \quad B_i \qquad \text{for } i = 1, \ldots, n$$
$$0 < P^*(A_i) = P^*(B_i) \qquad \text{for } i = 1, \ldots, n, \quad \text{and} \quad \sum_i P^*(A_i) = 1,$$

then $f \sim g$. Under these hypotheses $S/\bigcup A_i$ and $S/\bigcup B_i$ are null events (Exercise 17). Hence, with $f' = f$ on $\bigcup A_i$, $f' = x_1$ on $S/\bigcup A_i$, $g' = g$ on $\bigcup B_i$, and $g' = x_1$ on $S/\bigcup B_i$, $f' \sim f$ and $g' \sim g$ so that $f \sim g \Leftrightarrow f' \sim g'$. Thus it will suffice to prove that $f \sim g$ when $\{A_1, \ldots, A_n\}$ and $\{B_1, \ldots, B_n\}$ are partitions of S.

$f \sim g$ if $n = 1$. Using induction on $n > 1$ we shall "eliminate" x_n. Thus, assume the theorem is true for $n - 1$, and with $n > 1$ let

$$A = A_n \cap B_n^c \quad \text{and} \quad B = B_n \cap A_n^c$$

so that $A \cap B = \varnothing$ and $P^*(A) = P^*(B)$, the latter by $P^*(A_n \cap B_n^c) + P^*(A_n \cap B_n) = P^*(A_n) = P^*(B_n) = P^*(B_n \cap A_n^c) + P^*(B_n \cap A_n)$. Let $D_i = B \cap A_i$ for $i = 1, \ldots, n$ so that $\{D_1, \ldots, D_{n-1}\}$ is a partition of B. Then, by (14.3), there is a partition $\{C_1, \ldots, C_{n-1}\}$ of A for which

$$P^*(C_i) = P^*(D_i) \qquad i = 1, \ldots, n-1. \tag{14.9}$$

Let $f_0 = f$ and define $f_1, \ldots, f_{n-1}$ recursively thus: $f_i = f_{i-1}$ on $(C_i \cup D_i)^c$, $f_i = x_n$ on D_i, $f_i = x_i$ on C_i. Figure 14.1 illustrates this along with g. By (14.9) and Lemma 14.2, $f_i \sim f_{i-1}$ given $C_i \cup D_i$ for $i = 1, \ldots, n-1$. Since $f_i = f_{i-1}$ on $(C_i \cup D_i)^c$, $f_i \sim f_{i-1}$ given $(C_i \cup D_i)^c$. Therefore, by Lemma 14.1, $f_i \sim f_{i-1}$ for $i = 1, \ldots, n-1$ so that $f \sim f_{n-1}$.

						B_n						
	$A = A_n \cap B_n^c$					$B = A_n^c \cap B_n$						
	C_1	C_2	C_3	$\cdots$	C_{n-1}	D_1	D_2	D_3	$\cdots$	D_{n-1}	$A_n \cap B_n$	$A_n^c \cap B_n^c$
$= f_0$	x_n	x_n	x_n	$\cdots$	x_n	x_1	x_2	x_3	$\cdots$	x_{n-1}	x_n	f
f_1	x_1	x_n	x_n	$\cdots$	x_n	x_n	x_2	x_3	$\cdots$	x_{n-1}	x_n	f
f_2	x_1	x_2	x_n	$\cdots$	x_n	x_n	x_n	x_3	$\cdots$	x_{n-1}	x_n	f
$\cdot$												$\cdot$
$\cdot$												$\cdot$
$\cdot$												$\cdot$
f_{n-1}	x_1	x_2	x_3	$\cdots$	x_{n-1}	x_n	x_n	x_n	$\cdots$	x_n	x_n	f
g	g	g	g	$\cdots$	g	x_n	x_n	x_n	$\cdots$	x_n	x_n	g

Figure 14.1

It remains to show that $f_{n-1} \sim g$. With $B_n = B \cup (A_n \cap B_n)$ let

$$\begin{aligned} f' &= f_{n-1} \quad \text{on} \quad B_n^c & \qquad g' &= g \quad \text{on} \quad B_n^c \\ &= x_{n-1} \quad \text{on} \quad B_n & &= x_{n-1} \quad \text{on} \quad B_n. \end{aligned}$$

Then, as shown by Figure 14.1, the only consequences that can occur with f' and g' are $x_1, \ldots, x_{n-1}$. x_n has been eliminated. By (14.9) and Figure 14.1, $P^*\{f_{n-1}(s) = x_i\} = P^*\{f(s) = x_i\}$. Hence

$$P^*\{f' = x_i\} = P^*\{g' = x_i\} = P^*(B_i) \qquad \text{for } i = 1, \ldots, n-2,$$
$$P^*\{f' = x_{n-1}\} = P^*\{g' = x_{n-1}\} = P^*(B_{n-1}) + P^*(B_n)$$

which fits our initial format with n replaced by $n - 1$. Thus $f' \sim g'$ by the induction hypothesis. Then, since $f' \sim g'$ given B_n, Lemma 14.1 requires $f' \sim g'$ given B_n^c. Then, since $f_{n-1} = f'$ and $g = g'$ on B_n^c, $f_{n-1} \sim g$ given B_n^c. Also, since $f_{n-1} = g$ on B_n, $f_{n-1} \sim g$ given B_n. Hence $f_{n-1} \sim g$ by Lemma 14.1. ◆

The Axioms of Chapter 8

Defining $<$ on $\mathcal{P}_s$ by

$$P < Q \Leftrightarrow f < g \qquad \text{whenever } P_f = P \text{ and } P_g = Q, \tag{14.10}$$

$P1$ and Theorem 14.3 imply that $<$ on $\mathcal{P}_s$ is a weak order. The second and third conditions of Theorem 8.2 follow from the next two lemmas.

LEMMA 14.3. $(P, Q, R \in \mathcal{P}_s,\ 0 < \alpha < 1,\ P1\text{–}P6) \Rightarrow (P < Q \Leftrightarrow \alpha P + (1 - \alpha)R < \alpha Q + (1 - \alpha)R)$.

LEMMA 14.4. $(P, Q \in \mathcal{P}_s, f \in F, P < Q, P \leqslant f \leqslant Q, P1\text{–}P6) \Rightarrow$ *there is one and only one* $\alpha \in [0, 1]$ *such that* $f \sim \alpha P + (1 - \alpha)Q$.

Of course $f < P \Leftrightarrow f < g$ when $P_g = P$, with similar definitions for $f \preceq P$, $f \sim P, \ldots$. Theorem 14.3 guarantees no ambiguity here as long as $P \in \mathcal{P}_s$. The fact that Lemma 14.4 holds for any $f \in F$ will be used in the next section.

Proof of Lemma 14.3. Throughout this proof and the proof of Lemma 14.4 *we shall take* $\{x_1, \ldots, x_n\} = \{x : P(x) > 0\}$ *with* $P(x_i) = \alpha_i$, $\{y_1, \ldots, y_m\} = \{x : Q(x) > 0\}$ *with* $Q(y_j) = \beta_j$, *so that* $\sum \alpha_i = \sum \beta_j = 1$, *and let* w *be a most preferred consequence in* $\{x_1, \ldots, x_n, y_1, \ldots, y_m\}$. $A(\alpha)$ *will denote an event in* S *for which* $P^* = \alpha$. *Equation* (14.3) *will be used freely to construct events with various probabilities.*

For Lemma 14.3 we shall consider $f < g$ given $D(\alpha)$ with $D(\alpha > 0) \subseteq S$, $P^*_D(f = x_i) = \alpha_i\ (i = 1, \ldots, n)$ and $P^*_D(g = y_j) = \beta_j\ (j = 1, \ldots, m)$. In view of Theorem 14.3 and the first paragraph of its proof, $f < g$ given $D(1) \Leftrightarrow P < Q$, and $f < g$ given $D(\alpha) \Leftrightarrow \alpha P + (1 - \alpha)R < \alpha Q + (1 - \alpha)R$. (When $\alpha < 1$ let $f = g$ on $D(\alpha)^c$ with probabilities on $D(\alpha)^c$ equal to $(1 - \alpha)$ times the positive $R(x)$.) To prove the lemma we shall show that if $f < g$ given $D(\alpha)$ for one $\alpha \in (0, 1]$ then $f < g$ given $D(\alpha)$ for every $\alpha \in (0, 1]$.

Thus suppose that $f < g$ given $D(\gamma)$. Then, by considering n part uniform partitions of $D(\gamma)$, it follows from Lemma 14.1 that $f < g$ given $D(\gamma/n)$ for every positive integer n. Moreover, $f < g$ given $D(r\gamma)$ for every rational number $r \in (0, 1/\gamma]$.

Let $0 < \beta < 1$ be such that $f < g$ given $D(\beta)$. Let

$$f^* = f \quad \text{and} \quad g^* = g \quad \text{on} \quad D(\beta), \qquad f^* = g^* = w \quad \text{on} \quad D(\beta)^c$$

so that $f^* < g^*$. Then, using *P6* m times (once for each y_j) and Lemma 14.1 if necessary (so as not to exhaust all of $1 - \beta$ before the m uses of *P6* are completed), we obtain g'' with $f^* < g''$ and

$$\begin{aligned} g'' = g^* = g \quad &\text{on} \quad D(\beta) \\ = y_j \quad &\text{on} \quad C_j(\lambda_j);\ \lambda_j > 0,\ C_j(\lambda_j) \subseteq D(\beta)^c,\ C_j \cap C_k = \varnothing, \\ &\qquad\qquad\qquad\qquad\qquad\qquad (j = 1, \ldots, m) \\ = w \quad &\text{on} \quad \left(D(\beta) \cup \left(\bigcup_{j=1}^{m} C_j(\lambda_j)\right)\right)^c. \end{aligned}$$

Taking $C_j(\delta\beta_j) \subseteq C_j(\lambda_j)$ for $j = 1, \ldots, m$ with $\delta > 0$, let $g^0 = g''$ except that $g^0 = w$ on $C_j(\lambda_j)/C_j(\delta\beta_j)$. Since $y_j \preceq w$, Lemma 14.1 implies that $g'' \preceq g^0$, with

$$\begin{aligned} g^0 = g \quad &\text{on} \quad D(\beta + \delta) = D(\beta) \cup \left(\bigcup_{j=1}^{m} C_j(\delta\beta_j)\right) \\ = w \quad &\text{on} \quad D(\beta + \delta)^c. \end{aligned}$$

Also take $f^0 = f^*$ except that $f^0 = x_i$ on $E_i(\delta\alpha_i)$ where the E_i form a partition of $\bigcup_{j=1}^{m} C_j(\delta\beta_j)$. Since $x_i \preccurlyeq w, f^0 \preccurlyeq f^*$ by Lemma 14.1 with

$$\begin{aligned} f^0 &= f \quad \text{on} \quad D(\beta + \delta) \\ &= w \quad \text{on} \quad D(\beta + \delta)^c. \end{aligned}$$

Then $f^0 < g^0$ since $f^0 \preccurlyeq f^* < g'' \preccurlyeq g^0$, and hence $f < g$ given $D(\beta + \delta)$. Since this holds for all δ in some interval $(0. t]$, it follows from this and the preceding paragraph that $f < g$ given $D(\alpha)$ for all $\alpha \in (0, 1)$. Also, since $f < g$ given $D(1/2)$, Lemma 14.1 gives $f < g$ given $D(1)$. ◆

Proof of Lemma 14.4. As in the proof of $C1$ of Theorem 8.3, $(P, Q \in \mathcal{T}_s, P < Q, 0 \leq \alpha < \beta \leq 1) \Rightarrow \beta P + (1 - \beta)Q < \alpha P + (1 - \alpha)Q$ follows readily from Lemma 14.3. Thus, under the hypotheses of Lemma 14.4 there is one and only one $\alpha \in [0, 1]$ such that

$$\beta P + (1 - \beta)Q < f \qquad \text{if } \beta > \alpha \tag{14.11}$$

$$f < \beta P + (1 - \beta)Q \qquad \text{if } \beta < \alpha. \tag{14.12}$$

Clearly, only α can satisfy $f \sim \alpha P + (1 - \alpha)Q$.

Suppose then that $\alpha P + (1 - \alpha)Q < f$. This requires $\alpha > 0$. Let

$$\begin{aligned} g &= x_i \quad \text{on} \quad D(\alpha a_i) & i &= 1, \dots, n \\ &= y_j \quad \text{on} \quad D((1 - \alpha)\beta_j) & j &= 1, \dots, m \end{aligned}$$

where $\{D(\alpha\alpha_1), \dots, D((1 - \alpha)\beta_m)\}$ is a partition of S. Then $P_g = \alpha P + (1 - \alpha)Q$. Hence $g < f$ by Theorem 14.3 and $P1$. Then by repeated uses of $P6$ obtain $g' < f$ where $g' = g$ except that $g' = w$ on $C_i(\gamma_i > 0) \subseteq D(\alpha\alpha_i)$ for $i = 1, \dots, n$. With $\beta < \alpha$ and $\alpha - \beta$ small, take $C_i(\gamma_i' > 0) \subseteq C_i(\gamma_i)$ with $\gamma_i' = (\alpha - \beta)\alpha_i$ and let $g^0 = g'$ except that $g^0 = x_i$ on $C_i(\gamma_i)/C_i(\gamma_i')$. By Lemma 14.1, $g^0 \preccurlyeq g'$ with

$$\begin{aligned} g^0 &= x_i \quad \text{on} \quad D(\beta\alpha_i) \subset D(\alpha\alpha_i) & i &= 1, \dots, n \\ &= y_j \quad \text{on} \quad D((1 - \alpha)\beta_j) & j &= 1, \dots, m \\ &= w \quad \text{on} \quad D(\alpha - \beta). \end{aligned}$$

Now change g^0 to h by partitioning $D(\alpha - \beta)$ into $\{D(\beta_1(\alpha - \beta)), \dots, D(\beta_m(\alpha - \beta))\}$ and replacing w on $D(\beta_j(\alpha - \beta))$ by y_j. By Lemma 14.1, $h \preccurlyeq g^0$, so that by transitivity $h < f$. But $P_h = \beta P + (1 - \beta)Q$ by construction and since $\beta < \alpha$ we have obtained a contradiction to (14.12). Hence $\alpha P + (1 - \alpha)Q < f$ is false. Similarly, $f < \alpha P + (1 - \alpha)Q$ is false for this leads to a contradiction of (14.11). Hence $f \sim \alpha P + (1 - \alpha)Q$. ◆

In view of the results of this section and those of Chapter 8 we can state the following theorem, in which P^* is as given by (14.2) and (14.3) through (14.1).

THEOREM 14.4. *P1–P6 imply that there is a real-valued function u on X such that*

$$f < g \Leftrightarrow E[u(f(s)), P^*] < E[u(g(s)), P^*], \qquad \textit{for all } P_f, P_g \in \mathcal{T}_s, \quad (14.13)$$

and when u satisfies this representation it is unique up to a positive linear transformation.

In the rest of this chapter, u is assumed to satisfy (14.13).

14.5 UTILITIES ARE BOUNDED

In proving that u on X is bounded, we shall use the following lemma, whose proof follows easily from $P7$.

LEMMA 14.5. *($P1$, $P2$, $P7$, $x < f$ given A and $x < g$ given A for every $x \in X$) $\Rightarrow f \sim g$ given A. ($P1$, $P2$, $P7$, $f < x$ given A and $g < x$ given A for every $x \in X$) $\Rightarrow f \sim g$ given A.*

In the proof of the following theorem sup $T = \infty$ means that T is a set of real numbers and, if $c \in \text{Re}$, $t > c$ for some $t \in T$. sup $T = \infty$ means that T is unbounded above. In addition, when f is such that $P^*\{u(f(s)) > d\} = 1$ for some number d, $E(u, P_f) = \infty$ means that $\sup\{E[\inf\{u(f(s)), c\}, P^*] : c \in \text{Re}\} = \infty$.

THEOREM 14.5. *($P1$–$P7$) $\Rightarrow$ u on X is bounded.*

Proof. Let $P1$–$P7$ hold and suppose that u on X is unbounded above. Using (14.3) construct a sequence $B_1, B_2, \ldots$ of disjoint events in S with $P^*(B_n) = 2^{-n}$ for $n = 1, 2, \ldots$. If $\bigcup_{n=1}^{\infty} B_n$ does not exhaust S, add $S/\bigcup B_n$ to B_1. Take $u(x_n) \geq 2^n$ for each n and let

$$f = x_n \quad \text{on} \quad B_n, \qquad n = 1, 2, \ldots,$$

so that $E[u(f(s)), P^*] = \infty$ since

$$E[\inf\{u(f(s)), 2^n\}, P^*] \geq \sum_{i=1}^{n} P^*(B_i)u(x_i) \geq \sum_{i=1}^{n} 2^{-i}2^i = n$$

for $n = 1, 2, \ldots$. Let x be any consequence. Then, for some $y \in \{x_1, x_2, \ldots\}$,

$$u(x) < E[\inf\{u(f(s)), u(y)\}, P^*]. \qquad (14.14)$$

Let $f' = f$ on $\{s : f(s) \preccurlyeq y\}$ and $f' = y$ on $\{s : y \prec f(s)\}$. Then $P_{f'} \in \mathcal{T}_s$ and $E[u(f'(s)), P^*] = E[\inf\{u(f(s)), u(y)\}, P^*]$ so that, by Theorem 14.4 and (14.14), $x < f'$. But $f' \preccurlyeq f$ by Lemma 14.1 since, by $P7$, $f' \preccurlyeq f$ given $\{s : y \prec f(s)\}$. Hence $x < f$. Therefore $x \prec f$ for every x.

Next, let z be such that $u(x_1) < u(z)$. Let $g = z$ on B_1 and $g = f$ on B_1^c. As in the preceding paragraph, $x \prec g$ for every x, so that $f \sim g$ by Lemma 14.5. But $f \prec g$ given B_1 since $x_1 \prec z$ and $P^*(B_1) > 0$, and $f \sim g$ given B_1^c since $f = g$ on B_1^c. Hence $f \prec g$ by Lemma 14.1, a contradiction. Hence u is bounded above. A symmetric proof shows that u is bounded below. ◆

14.6 UTILITY FOR ALL ACTS

To establish $f \prec g \Leftrightarrow E(u, P_f) < E(u, P_g)$ for all acts we shall first prove two lemmas.

With $P \in \mathcal{P}_s$, $g \doteq P$ on A means that $P^*\{s \in A \text{ and } g(s) = x\} = P^*(A)P(x)$ for all $x \in X$. We define $f \prec P$ given $A \Leftrightarrow f \prec g$ given A for every $g \doteq P$ on A. $P \prec f$ given A is similarly defined, and $f \sim P$ given A and $f \preccurlyeq P$ given A are defined in the usual way. Note that, by Theorem 14.3, if $f \prec g$ given A for one $g \doteq P \in \mathcal{P}_s$ on A, then $f \prec h$ given A for every $h \doteq P$ on A. If A is null, $g \doteq P$ on A for every g.

LEMMA 14.6. *($P1$–$P7$, $A \neq \varnothing$, $f \preccurlyeq x$ given A, $u(f(s)) < c$ for all $s \in A$) $\Rightarrow$ there is a $P \in \mathcal{P}_s$ for which $f \preccurlyeq P$ given A and $E(u, P) \leq c$. ($P1$–$P7$, $A \neq \varnothing$, $x \preccurlyeq f$ given A, $c < u(f(s))$ for all $s \in A$) $\Rightarrow$ there is a $P \in \mathcal{P}_s$ for which $P \preccurlyeq f$ given A and $c \leq E(u, P)$.*

LEMMA 14.7. *($P1$–$P7$, $\{B_1, \ldots, B_n\}$ is a partition of S, $u(f(s)) < c_i$ for all $s \in B_i$ $(i = 1, \ldots, n)$, $P \in \mathcal{P}_s$, $P \preccurlyeq f$) $\Rightarrow E(u, P) \leq \sum_{i=1}^n P^*(B_i)c_i$. ($P1$–$P7$, $\{B_1, \ldots, B_n\}$ is a partition of S, $c_i < u(f(s))$ for all $s \in B_i$ $(i = 1, \ldots, n)$, $P \in \mathcal{P}_s$, $f \preccurlyeq P$) $\Rightarrow \sum_{i=1}^n P^*(B_i)c_i \leq E(u, P)$.*

It will suffice to prove the first part of each lemma. In each proof the hypotheses of the first part are assumed to hold.

Proof of Lemma 14.6. If $u(x) \leq c$ let $P(x) = 1$. Then $f \preccurlyeq P$ given A by hypothesis and $E(u, P) = u(x) \leq c$. Henceforth suppose that $c < u(x)$. Let y be any consequence for which $u(y) \leq c$, as assured by $A \neq \varnothing$ and $u(f(s)) < c$ for all $s \in A$. Let P be the unique combination of x and y for which $E(u, P) = c$. If A is null then $f \preccurlyeq P$ given A and the proof is complete. Henceforth assume that $P^*(A) > 0$.

Fix $t \in A$. Let $g \doteq P$ on A and $g = f(t)$ on A^c. Since $u(f(t)) < c$,

$$u(f(t)) = P^*(A)u(f(t)) + P^*(A^c)u(f(t))$$
$$< P^*(A)E(u, P) + P^*(A^c)u(f(t)) = E[u(g(s)), P^*]$$

so that, by Theorem 14.4, $f(t) \prec g$. Hence $f(t) \prec g$ given A. Since this holds

for each $t \in A$, $P7$ implies that $f \preccurlyeq g$ given A. Since $g \doteq P$ on $A, f \preccurlyeq P$ given A. ◆

Proof of Lemma 14.7. Suppose the conclusion is false for some f and P so that $\sum P^*(B_i)c_i < E(u, P)$. Since this can't hold if P is confined with probability 1 to worst consequences, it follows that there is a $Q \in \mathcal{P}_s$ for which $\sum P^*(B_i)c_i < E(u, Q)$ and $Q < P \preccurlyeq f$. Hence, if the lemma is true when its $P \preccurlyeq f$ hypothesis is replaced by $P < f$ then the original lemma must be true. Thus, it will suffice to show that if $P1$–$P7$ hold, if $\{B_1, \ldots, B_n\}$ is a partition of S and if

1. $u(f(s)) < c_i$ for all $s \in B_i$, $i = 1, \ldots, n$, and
2. $P \in \mathcal{P}_s$ and $P < f$,

then $E(u, P) \leq \sum_{i=1}^n P^*(B_i)c_i$.

To prove this we show first that f can be modified, if necessary, so that (1) and (2) hold for the modified f and, for each i, there is a y_i such that modified $f \preccurlyeq y_i$ given B_i. If there is a y_i such that $f \preccurlyeq y_i$ given B_i, we cease to worry about this i. On the other hand, suppose $x < f$ given B_i for every $x \in X$. Then B_i can't be null so that $P^*(B_i) > 0$. For this B_i take $y < z$ and $u(y) < c_i$. With $P < f$ by (2), it follows from $P6$ that there is a non-null $A \subseteq B_i$ for which $P < f'$ when $f' = f$ except on A where $f' = y$. Let $f^* = f$ except on A where $f^* = z$. Since $y < z$, $f' < f^*$ given A. Hence $f' < f^*$ given B_i by Lemma 14.1. It cannot be true that $x < f'$ given B_i for every $x \in X$ for otherwise, by Lemma 14.5, $f' \sim f^*$ given B_i, a contradiction. Hence there is a $y_i \in X$ such that $f' \preccurlyeq y_i$ given B_i. Since (1) and (2) hold for f' we see that, by considering each i, we obtain an act g that satisfies

1. $u(g(s)) < c_i$ for all $s \in B_i$, $i = 1, \ldots, n$,
2. $P \in \mathcal{P}_s$ and $P < g$,
3. There is a $y_i \in X$ for which $g \preccurlyeq y_i$ given B_i, $i = 1, \ldots, n$.

Given such a g, Lemma 14.6 implies that, for each i, there is a $Q_i \in \mathcal{P}_s$ such that $g \preccurlyeq Q_i$ given B_i and $E(u, Q_i) \leq c_i$. Let $h \doteq Q_i$ on B_i for $i = 1, \ldots, n$. Then, by Lemma 14.1, $g \preccurlyeq h$ so that $P < h$. Since $P_h = \sum P^*(B_i)Q_i$, Theorem 14.4 implies that $E(u, P) < E[u(h(s)), P^*]$. Since $E[u(h(s)), P^*] = \sum P^*(B_i)E(u, Q_i) \leq \sum P^*(B_i)c_i$, $E(u, P) < \sum P^*(B_i)c_i$. ◆

Expected Utility for All Acts

THEOREM 14.6. $P1$–$P7 \Rightarrow$ (14.4).

Proof. By an appropriate positive linear transformation of u, we use $P5$ and Theorem 14.5 to specify

$$\inf\{u(x) : x \in X\} = 0, \qquad \sup\{u(x) : x \in X\} = 1.$$

Each act in F falls into exactly one of the following classes:

1. f is *big* $\leftrightarrow x < f$ for every $x \in X$,
2. f is *little* $\Leftrightarrow f < x$ for every $x \in X$,
3. f is *normal* $\Leftrightarrow x \leqslant f \leqslant y$ for some $x, y \in X$.

Suppose first that f is normal. Lemma 14.4 guarantees that there is a $P \in \mathfrak{T}_s$ such that $P \sim f$. Divide S into $A_1 = \{s:0 \leq u(f(s)) \leq 1/n\}$, $A_i = \{s:(i-1)/n < u(f(s)) \leq i/n\}$ for $i = 2, \ldots, n$. Some of the A_i may be empty. By the definition of expectation (Definition 10.12, Exercise 10.16), $\sum_i P^*(A_i)(i-1)/n \leq E[u(f(s)), P^*] \leq \sum_i P^*(A_i)i/n$. Also, by Lemma 14.7, $\sum_i P^*(A_i)(i - 1 - \epsilon)/n \leq E(u, P) \leq \sum_i P^*(A_i)(i + \epsilon)/n$ for any $\epsilon > 0$. Letting n get large it follows that

$$E[u(f(s)), P^*] = E(u, P) \qquad \text{when } f \sim P, P \in \mathfrak{T}_s. \tag{14.15}$$

Suppose next that f is big. By Lemma 14.5, all big acts are indifferent. We shall prove that

$$f \textit{ is big} \Rightarrow u(x) < 1 \textit{ for all } x, P^*\{u(f(s)) \geq 1 - \epsilon\} = 1$$
$$\textit{for } \epsilon > 0, E[u(f(s)), P^*] = 1.$$

With f big suppose first that $u(w) = 1$ for $w \in X$. Take $x < w$, using *P5*. Let $A = \{s:u(f(s)) < 1\}$, $A^c = \{s:u(f(s)) = 1\}$. Then, using *P7* if A is not null, as in the final part of the proof of Lemma 14.6, it follows that $f \leqslant w$ given A. [Suppose that $w < f$ given A^c (requiring A^c to be non-null). Then, by *P6*, there is a non-null $B \subseteq A^c$ with $w < f'$ given A^c and $f' = f$ except on B where $f' = x$. Let $f'' = f$ except on B where $f'' = w$. Then $f' < f''$ given A^c by Lemma 14.1. But then, using Lemma 14.5, $f' \sim f''$ given A^c, a contradiction.] Hence $f \leqslant w$ given A^c so that $f \leqslant w$ by Lemma 14.1. But $f \leqslant w$ contradicts f's bigness. Hence f is big $\Rightarrow u(x) < 1$ for all $x \in X$.

Suppose next that for big f there is an $\epsilon > 0$ for which $P^*\{u(f(s)) \geq 1 - \epsilon\} < 1$. Then, with $A = \{s:u(f(s)) < 1 - \epsilon\}$, $P^*(A) > 0$. It follows from the preceding paragraph that we can select $y, z \in X$ so that

$$1 - \epsilon < u(y) < u(z) < 1.$$

Let $f'' = f' = f$ except on A where $f' = y$ and $f'' = z$. Then, since $u(f(s)) < u(y)$ for all $s \in A$, $f \leqslant y$ given A. This leads to $f \leqslant f' < f''$. But since f is big f'' is then big also and hence $f \sim f''$ by Lemma 14.5, a contradiction. Therefore $P^*\{u(f(s)) \geq 1 - \epsilon\} = 1$ for every $\epsilon > 0$. Therefore $E[u(f(s)), P^*] \geq 1 - \epsilon$ for every $\epsilon > 0$ and, since E can't exceed 1 (Exercise 10.22*a*), $E[u(f(s)), P^*] = 1$.

By a symmetric proof for little acts it follows that

$$f \textit{ is little} \Rightarrow 0 < u(x) \textit{ for all } x, P^*\{u(f(s)) \leq \epsilon\} = 1$$
$$\textit{for } \epsilon > 0, E[u(f(s)), P^*] = 0,$$

and, by Lemma 14.5, all little acts are indifferent to each other.

(14.4) follows readily from Theorem 14.4, Lemma 14.5, the fact that every normal act is indifferent to some $P \in \mathcal{P}_s$, and from (14.15) and the implications for big and little acts. ◆

14.7 SUMMARY

Savage's axioms for expected utility apply $<$ to the set F of all functions on S to X (states to consequences). When $<^*$ (is less probable than) is defined on the basis of $<$ is an appropriate way, his first six axioms imply that there is a probability measure P^* on S that satisfies $A <^* B \Leftrightarrow P^*(A) < P^*(B)$, for all A, $B \subseteq S$, and, when this holds, $(B \subseteq S,\ 0 \leq \rho \leq 1) \Rightarrow P^*(C) = \rho P^*(B)$ for some $C \subseteq B$, and P^* is unique. This latter property implies that the set $\{P_f : f \in F\}$ of probability measures on X induced by P^* on S includes the set $\mathcal{P}_s$ of all simple measures on X. By showing that axioms similar to those of Chapter 8 follow for $<$ on $\mathcal{P}_s$, we obtain an expected-utility representation for $\mathcal{P}_s$, or for the set of simple acts. Savage's seventh axiom then implies that the utility function u on X is bounded and that the expected-utility representation $f < g \Leftrightarrow E[u(f(s)), P^*] < E[u(g(s)), P^*]$, or equivalently $f < g \Leftrightarrow E(u, P_f) < E(u, P_g)$, holds for all acts.

Savage's book (1954) contains an excellent section on "Historical and critical comments on utility" (pp. 91–104) that should be studied by everyone interested in utility.

INDEX TO EXERCISES

1–2. Probability axioms for finite sets. 3. A/B. 4. $C1$. 5. Qualitative probability implications. 6. $F5$. 7–9. Uniform partitions, almost agreeing measures. 10–14. Fine and tight qualitative probabilities. 15. $<$ given A. 16. Failure of $P2$. 17. A is null $\Leftrightarrow A \sim^* \varnothing$. 18. Discrete measures in $\mathcal{P}$. 19. Conditional probability. 20. $P1$–$P6$ hold, $P7$ fails. 21. A variant of $P7$. 22. $P1$–$P7$ do not imply: $P^*\{f(s) < g(s)\} = 1 \Rightarrow f < g$.

Exercises

1. Kraft, Pratt, and Seidenberg (1959), Scott (1964). Use the Theorem of The Alternative (Theorem 4.2) to prove the following theorem. *Suppose that S is finite. Then there is a binary relation $<^*$ on the set of all subsets of S that satisfies* (14.2) *if and only if, for all $s \in S$, all $A_1, \ldots, B_m \subseteq S$ and all $m \geq 2$:*

1. *not* $\{s\} <^* \varnothing$,
2. $\varnothing <^* S$,
3. $(\sum_1^m A_j = \sum_1^m B_j,\ A_j <^* B_j$ *or* $A_j \sim^* B_j$ *for each* $j < m) \Rightarrow$ *not* $A_m <^* B_m$.

In (3), $A \sim^* B \Leftrightarrow$ (not $A \prec^* B$, not $B \prec^* A$), and $\sum_1^m A_i = \sum_1^m B_i \Leftrightarrow$ for each s, the number of A_i that contain s equals the number of B_i that contain s.

2. (*Continuation.*) Kraft, Pratt, and Seidenberg (1959). Let $S = \{p, q, r, s, t\}$ and denote a subset of S such as $\{p, q, t\}$ by pqt. Let $\prec^*$ on the set of all events be given by

$$\begin{aligned} &\varnothing \prec^* p \prec^* q \prec^* r \prec^* pq \prec^* pr \prec^* s \prec^* ps \\ &\prec^* qr \prec^* t \prec^* pqr \prec^* qs \prec^* rs \prec^* pt \prec^* pqs \prec^* qt \\ &\prec^* prs \prec^* rt \prec^* qrs \prec^* pqt \prec^* prt \prec^* st \prec^* pqrs \prec^* pst \\ &\prec^* qrt \prec^* pqrt \prec^* qst \prec^* rst \prec^* pqst \prec^* prst \prec^* qrst \prec^* pqrst, \end{aligned}$$

in which the order of the last two rows is the order of the complements the first two rows in reverse. Clearly, $F1$, $F2$, and $F3$ of Theorem 14.2 hold.

a. Show that $F4$ holds.
b. Show that condition (3) in Exercise 1 fails.

3. For any $A, B \subseteq S$ verify that

a. $(A/B) \cap (A \cap B) = \varnothing$ and $(A/B) \cup (A \cap B) = A$,
b. $A \cup (B/A) = B \cup (A/B) = A \cup B$,
c. $(A/B) \cap (B/A) = \varnothing$,
d. $(A/B) \cup (B/A) = (A \cup B)/(A \cap B)$,
e. $(A/B) \cup (B/A) \cup (A \cap B) = A \cup B$,
f. $(A/B) \cup (B/C) = (A/C) \cup ((A \cap C)/B) \cup (B/(A \cap C))$, with A/C, $(A \cap C)/B$, and $B/(A \cap C)$ mutually disjoint.

4. Prove $C1$ of Section 14.2.

5. Let $\prec^*$ satisfy $F1$–$F4$. Verify

a. $A \prec^* B \Leftrightarrow A/B \prec^* B/A$,
b. $A \prec^* B \Leftrightarrow B^c \prec^* A^c$,
c. $(A \prec^* A^c, B \prec^* B^c) \Rightarrow A \prec^* B^c$,
d. $S \preccurlyeq^* B \Rightarrow B \sim^* S$ and $A \cap B \sim^* A$,
e. $(A \sim^* B, C \sim^* D, A \cup B \sim^* C \cup D, A \cap B = C \cap D = \varnothing) \Rightarrow A \sim^* C$.

6. Without using $C5$–$C9$, prove that $(F1$–$F5, \varnothing \prec^* A, \varnothing \prec^* B) \Rightarrow \varnothing \prec^* C \prec^* A$ for some $C \subseteq B$.

7. (*Continuation.*) Let $F6$ be: *If* $\varnothing \prec^* A$ *then A can be partitioned into B and C with* $B \sim^* C$, *as in C8. On examining the proof of* (14.2) *through step* 6, *argue that F1–F4 and F6 imply that there is a unique probability measure* P^* *on the set of events that satisfies*

$$A \preccurlyeq^* B \Rightarrow P^*(A) \le P^*(B), \qquad \textit{for all} \quad A, B \subseteq S. \tag{14.16}$$

8. (*Continuation.*) Show that the proof of (14.3) holds for the situation of the preceding exercise, so that $F1$–$F4$ *and F6 imply* (14.3) *when* P^* *satisfies* (14.16).

9. Let $F7$ be: *For every position integer n there is an n part* u.p. *of S*. On examining the proof of (14.2) through step 6, argue that $F1$–$F4$ and $F7$ *imply that there is a unique probability measure* P^* *on the set of events that satisfies* (14.16). Also prove that $(F1$–$F4$, $F7) \Rightarrow$ (14.3).

10. Following Savage (pp. 36–37) we define the following terms for a qualitative probability $\prec^*$ on the events in S (that satisfies $F1$–$F4$):

$\prec^*$ is *fine* $\Leftrightarrow$ ($\varnothing \prec^* A \Rightarrow$ there is a finite partition of S each element of which is not more probable than A).

$\prec^*$ is *tight* $\Leftrightarrow A \sim^* B$ whenever $A \preceq^* B \cup C$ and $B \preceq^* A \cup D$ for all C and D that satisfy ($B \cap C = A \cap D = \varnothing$, $\varnothing \prec^* C$, $\varnothing \prec^* D$).

Given $F1$–$F4$, prove that $\prec^*$ is both fine and tight $\Leftrightarrow F5$ holds.

11. (*Continuation.*) Following Savage (p. 41), let $S_1 = [0, 1]$, $S_2 = [2, 3]$ and let P_i be a finitely additive probability measure on the set of all subsets of S_i ($i = 1, 2$) that agrees with Lebesgue measure [e.g., $P_1([a, b]) = b - a$ when $0 \le a \le b \le 1$] on the Lebesgue measurable subsets of S_i. Let $S = S_1 \cup S_2$ and, for any $A \subseteq S$ let $A_1 = A \cap S_1$ and $A_2 = A \cap S_2$. Define $\prec^*$ on the set of all subsets of S as follows: $A \prec^* B \Leftrightarrow P_1(A_1) < P_1(B_1)$ or $(P_1(A_1) = P_1(B_1), P_2(A_2) < P_2(B_2))$.

a. Verify that $\prec^*$ is a qualitative probability. ($F1$–$F4$ hold.)
b. Prove that $\prec^*$ is not fine. [Let $A = S_2$ and argue that any finite partition of S must contain a B for which $A \prec^* B$.]
c. Prove that $\prec^*$ is tight.
d. With $P^*(A) = P_1(A_1)$ for all $A \subseteq S$, does P^* satisfy (14.16)?

12. (*Continuation.*) Let S_1, S_2, P_1, P_2, and S be defined as in the preceding exercise, let $A_1 = A \cap S_1$ and $A_2 = A \cap S_2$ for any $A \subseteq S$, and define $A \prec^* B \Leftrightarrow P_1(A_1) + P_2(A_2) < P_1(B_1) + P_2(B_2)$ or $(P_1(A_1) + P_2(A_2) = P_1(B_1) + P_2(B_2), P_1(A_1) < P_1(B_1))$.

a. Show that $\prec^*$ is a qualitative probability.
b. Prove that $\prec^*$ is fine.
c. Prove that $\prec^*$ is not tight. [Let $A = S_1$, $B = S_2$ and show that if $\prec^*$ is tight then $A \sim^* B$. But, by definition, $B \prec^* A$.]
d. With $P^*(A) = \frac{1}{2}[P_1(A_1) + P_2(A_2)]$, does P^* satisfy (14.16)?

13. (*Continuation.*) Let S_1, S_2, P_1, and P_2 be as given in Exercise 11. Let $S_3 = [4, 5]$ and let P_3 be a finitely additive extension of Lebesgue measure on S_3. Take $S = S_1 \cup S_2 \cup S_3$, let $A_i = A \cap S_i$ for 1, 2, 3 and any $A \subseteq S$, and define $\prec^*$ by $A \prec^* B \Leftrightarrow P_1(A_1) < P_1(B_1)$ or $(P_1(A_1) = P_1(B_1),\ P_2(A_2) + P_3(A_3) < P_2(B_2) + P_3(B_3))$ or $(P_1(A_1) = P_1(B_1), P_2(A_2) + P_3(A_3) = P_2(B_2) + P_3(B_3), P_2(A_2) < P_2(B_2))$.

a. Verify that $\prec^*$ is a qualitative probability.
b. Show that $\prec^*$ is not fine.
c. Show that $\prec^*$ is not tight.
d. With $P^*(A) = P_1(A_1)$, does P^* satisfy (14.16)?

14. (*Continuation.*) In each of the three preceding exercises argue that, for each positive integer n, there is an n part uniform partition of S. It then follows from Exercise 9 that P^* as defined for each of the three preceding exercises is the only probability measure on S that satisfies (14.16). Then show that, in each of the three cases, there are $A, B \subseteq S$ for which $A \prec^* B$ and $P^*(A) = P^*(B)$, so that (14.2) cannot hold.

Note: In the remaining exercises F is the set of all functions on S to X.

15. Prove that $(P1, P2) \Rightarrow \prec$ given A is a weak order.

16. Savage (correspondence). Let $S = [0, 1]$ with P^* on S an extension of Lesbesgue measure on $[0, 1]$ so that, for example, $P^*([a, b]) = b - a$ when $0 \leq a \leq b \leq 1$. Let $X = [0, \infty)$ and take $u(x) = x$, so that F is the set of all nonnegative real functions on $[0, 1]$. Admitting the case of $E[u(f(s)), P^*] = E(f, P^*) = \infty$ (see Section 14.5), with $\infty = \infty$, take $f \prec g$ if and only if $E(f, P^*) < E(g, P^*)$.

a. Show that $P2$ fails in this situation, by considering four acts with $f = f' = 1$, $g = g' = 0$ on $A = [0, \frac{1}{2})$, and $f = g$ and $f' = g'$ on A^c with $E(f, P^*) = \infty$ and $E(f', P^*)$ finite.

b. Verify that $P1$ and $P3$–$P7$ hold.

17. Prove that if $P1$–$P5$ hold then A is null $\Leftrightarrow A \sim^* \varnothing$.

18. Verify that (14.3) implies that all discrete probability measures on X are in $\mathcal{P} = \{P_f : f \in F\}$ under (14.8).

19. Let P_A^* be the conditional probability measure of P^* given A when $P^*(A) > 0$, with $P_A^*(B) = P^*(A \cap B)/P^*(A)$ for all $B \subseteq S$. Verify that $P1$–$P7$ imply, for all $f, g \in F$ and $A \subseteq S$, that $f \precsim g$ given $A \Leftrightarrow P^*(A) = 0$, or $E[u(f(s)), P_A^*] \leq E[u(g(s)), P_A^*]$ when $P^*(A) > 0$.

20. Savage (p. 78). Let $S = \{1, 2, \ldots\}$, let $X = [0, 1)$, and let P^* be a diffuse measure on S with $P^*(s) = 0$ for all $s \in S$ and $P^*\{n + j, 2n + j, 3n + j, \ldots\} = 1/n$ for all $n > 0, j \geq 0$. Define $\prec$ on F by $f \prec g \Leftrightarrow w(f) < w(g)$ where

$$w(f) = E(f, P^*) + \inf\{P^*\{f(s) \geq 1 - \epsilon\} : \epsilon > 0\}.$$

a. Prove that if $\{A_1, \ldots, A_n\}$ is a partition of S and if $I = \{i : P^*(A_i) > 0\}$ then, with $P_{A_i}^*$ as defined in Exercise 19,

$$w(f) = \sum_I P^*(A_i)[E(f, P_{A_i}^*) + \inf\{P_{A_i}^*\{f(s) \geq 1 - \epsilon\} : \epsilon > 0\}].$$

b. Verify that $P1$–$P6$ hold.

c. Show that $P7$ is violated by f and g where $f = 0$ and $g = \frac{1}{2}$ on the odd integers and $f(n) = g(n) = n/(n + 1)$ for each even integer n.

21. (*Continuation.*) In Chapter 10 we saw that Axiom $A4b$ (Section 10.4) is not sufficient for the general expected-utility result when the measures in $\mathcal{P}$ are not all countably additive. In the context of the present chapter the correspondent of $A4b$ is $P7b$: *If $x \precsim g(s)$ given A for all $s \in A$ then $x \precsim g$ given A; if $g(s) \precsim x$ given A for all $s \in A$ then $g \precsim x$ given A.* Verify that $P7b$ holds for the example of the preceding exercise.

22. Modify the example of Exercise 20 to give a case where $P1$–$P7$ hold and where the following assertion is false: If $f(s) \prec g(s)$ for all $s \in A$ and A is not null, then $f \prec g$ given A.

ANSWERS TO SELECTED EXERCISES

2.1b. For each $i \in \{\ldots, -2, -1, 0, 1, 2, \ldots\}$, let $f(0) = 1$, $f(i) = 2i$ when $i > 0$ and $f(i) = -2i + 1$ when $i < 0$.

2.4a. $\sim$ is reflexive since $\preccurlyeq$ is reflexive. $\sim$ is symmetric from its definition. If $x \sim y$ and $y \sim z$ then $(x \preccurlyeq y, y \preccurlyeq x)$ and $(y \preccurlyeq z, z \preccurlyeq y)$, which by the transitivity of $\preccurlyeq$ yield $(x \preccurlyeq z, z \preccurlyeq x)$.

2.5. If $x \prec^t x$ then not $x \prec^t x$ by asymmetry, a contradiction. Hence $\prec^t$ is irreflexive. Suppose $x \prec^t y$, $y \prec^t z$. Then $x \prec x_1, x_1 \prec x_2, \ldots, x_m \prec y, y \prec y_1, y_1 \prec y_2, \ldots, y_n \prec z$, so that $x \prec^t z$.

2.7. $X = \{x, y, z\}$ with $x \prec y$, $y \prec z$, and $x \sim z$.

2.9. Define x and y as equivalent if and only if they are in the same element of the partition.

2.11. Suppose $\prec$ is transitive and $x \sim y \Leftrightarrow I(x) \cap I(y) \neq \varnothing$. Clearly, $\prec$ is irreflexive since $I(x) \cap I(x) \neq \varnothing$. Suppose $(x \prec y, z \prec w)$ so that $I(x) \cap I(y) = \varnothing$ and $I(z) \cap I(w) = \varnothing$. Then either $I(x) \cap I(w) = \varnothing$ in which case either $x \prec w$ or $w \prec x$ (and hence $z \prec y$ by transitivity); or $I(x) \cap I(z) = \varnothing$ in which case either $x \prec z$ (and hence $x \prec w$ by transitivity) or $z \prec x$ (and hence $z \prec y$); or $I(y) \cap I(z) = \varnothing \cdots$; or $I(y) \cap I(w) = \varnothing \cdots$.

2.14. $TE \Rightarrow$ Transitivity. Let TE hold. Suppose Transitivity fails with $y \in F(\{x, y\})$, $z \in F(\{y, z\})$, and $\{x\} = F(\{x, z\})$. Then $z \notin F(\{x, y, z\})$ by TE. If $y \in F(\{x, y, z\})$ then $z \notin F(\{y, z\})$ by TE, a contradiction. Hence $y \notin F(\{x, y, z\})$. Therefore $\{x\} = F(\{x, y, z\})$. Then, by TE, $y \notin F(\{x, y\})$, another contradiction. Therefore $F(\{x, y, z\}) = \varnothing$, which is false. Hence Transitivity must hold when TE holds.

2.15. $f(x_1, x_2) = x_1 + .5(x_1 + x_2 - 2)(x_1 + x_2 - 1)$ gives a one-to-one correspondence.

2.17. $u(x_1, x_2) = ax_1 + x_2$ with $a > 1$ will do.

2.21b. $(x, y) \in (A \cup B)' \Leftrightarrow (y, x) \in A$ or $(y, x) \in B$. $(x, y) \in A' \cup B' \Leftrightarrow (y, x) \in A$ or $(y, x) \in B$. $(x, y) \in (A \cap B)' \Leftrightarrow (y, x) \in A$ and $(y, x) \in B$.

3.1. If the subset is countable let it be enumerated as $0.x_{11}x_{21}x_{31}\cdots$, $0.x_{12}x_{22}x_{32}\cdots$, $0.x_{13}x_{23}x_{33}\cdots$, $\cdots$. Let $x_i \neq x_{ii}$ for all i, $x_i \in \{1, 2\}$. Is $0.x_1x_2x_3\cdots$ in the enumeration?

3.3. Let $U(x) = (|x|, x)$. Then $x \prec y$ if and only if $U(x) <^L U(y)$ where $(a, b) <^L (c, d)$ if and only if $a < c$ or $[a = c$ and $b < d]$.

3.6c. $(1, -2, 4, -3)$.

3.7b. 130.

3.8. Let X be the unit square with $(0, 1) \prec (1, 0)$. The set of all $\alpha(0, 1) + (1 - \alpha)(1, 0) = (1 - \alpha, \alpha)$ is the straight line segment from $(0, 1)$ to $(1, 0)$. Can you draw a valid indifference curve that intersects this segment at several places?

3.12. Yes. If the $x \in X$ are numbers, X is finite, and $\mathcal{T} = \{A \cap X : A \in \mathcal{U}\}$, then u on X is continuous in $\mathcal{T}$.

3.16. Given $c \in (u(x), u(y))$ suppose $c \neq u(z)$ for every $z \in X$. Let $Y = \{z : z \in X, u(z) < c\}$, $Z = \{z : z \in X, c < u(z)\}$. Y and Z are nonempty, disjoint, $Y \cup Z = X$, and since $\{b : b < c\}$ and $\{a : c < a\}$ are in $\mathcal{U}$ and u is continuous, $Y, Z \in \mathcal{T}$, contradicting the connectedness of $(X, \mathcal{T})$.

3.23. If X is not connected then it can be partitioned into nonempty, disjoint subsets Y and Z that are both in $\{A \cap X : A \in \mathcal{U}^n\}$. If X is convex and $y \in Y$, $z \in Z$, then $L(y, z) = \{\alpha y + (1 - \alpha)z : \alpha \in [0, 1]\}$ is in X and $(L(y, z), \{A \cap L(y, z) : A \in \mathcal{U}^n\})$ is connected. But by X not being connected we must conclude that $L(y, z) \cap Y$ and $L(y, z) \cap Z$ are nonempty, disjoint open sets in $\{A \cap L(y, z) : A \in \mathcal{U}^n\}$ that partition $L(y, z)$, so that $L(y, z)$ is not connected, a contradiction.

4.2. $4! = 24$. Eight are additive.

4.3. a, c, f, g.

4.4. $x_1x_2 + (x_1x_2)^2 = x_1x_2(1 + x_1x_2)$. With $a, b \geq 1$, $a \leq b \Leftrightarrow a(1 + a) \leq b(1 + b)$. See Exercise 3a.

4.5. Hint: Include in $(\quad) E_4 (\quad)$ the six elements whose utilities are duplicated $(8, 9, 15)$.

4.12. Yes.

4.15. For Theorem 4.2 let $C = (u_1(x_{11}), u_1(x_{12}), \ldots, u_n(x_{nr}), \sigma(x^1), \sigma(x^2), \ldots, \sigma(x^s))$.

4.17. Let $c = (u_1(x_{11}), \ldots, u_n(x_{nr}), 1)$.

5.1a. $u_1(0) = 0$, $u_1(1) = 2$, $u_2(r) = 2 - e^{-r}$ when $r \geq 0$ and $u_2(r) = e^r$ when $r < 0$ will do.

5.1b. Assume additive utilities exist, let $\alpha = u_2(1) - u_2(0) > 0$, $M = u_1(1) - u_1(0) > 0$, $\beta_i = u_1(1/i) - u_1(1/(1 + i))$ for $i = 1, 2, \ldots$ and show that $M > m\alpha$ for every positive integer m.

5.4. Suppose $m = 3$, $n = -2$. Then $3x - 2x = x + x + (x - x) - x = x + x + e - x = x + (x - x) = x + e = x$.

5.12. X.

5.13. $\Pi^* \mathcal{T}_i \subseteq \Pi \mathcal{T}_i$ since $\prod_{i=1}^n A_i \in \Pi \mathcal{T}_i$ when $A_i \in \mathcal{T}_i$ for each i. Suppose $A \neq \varnothing$, $A \in \Pi \mathcal{T}_i$. For $x \in A$ let $A_i(x) \in \mathcal{T}_i$ be such that $x_i \in A_i(x)$, $\Pi A_i \subseteq A$, $\Pi A_i \in \Pi^* \mathcal{T}_i$. Also, $\bigcup_{x \in A} (\Pi A_i(x)) = A$ so that $A \in \Pi^* \mathcal{T}_i$. Thus, $\Pi \mathcal{T}_i \subseteq \Pi^* \mathcal{T}_i$.

5.19. Suppose $U \in \mathcal{U}$. Then $U = \bigcup_{t \in T} A(t)$, where $A(t)$ is an open interval in Re for each $t \in T$, by Exercise 18. Therefore $f^{-1}(U) = \bigcup_{t \in T} f^{-1}(A(t))$ is in $\mathcal{T}$ when $f^{-1}(A(t))$ is in $\mathcal{T}$ for every $t \in T$.

5.23. $(x_1, y_1, z_1, x_2, y_2, z_2) = (2, 1, 3, 1.5, 5, 8)$ gives $u(x_1, x_2) < u(y_1, y_2)$, $u(y_1, z_2) < u(z_1, x_2)$, and $u(z_1, y_2) < u(x_1, z_2)$, which contradict $Q1$.

6.1b. $[x^1, \ldots, x^m, y^1, \ldots, y^m$ is a permutation of $z^1, \ldots, z^m, w^1, \ldots, w^m$, $(x^j, y^j) \prec (z^j, w^j)$ or $(x^j, y^j) \sim (z^j, w^j)$ for all $j < m] \Rightarrow$ not $(x^m, y^m) \prec (z^m, w^m)$. Alternatively, $[(x^1, y^1), \ldots, (x^m, y^m) E_m (z^1, w^1), \ldots, (z^m, w^m)$, $(x^j, y^j) \prec (z^j, w^j)$ or $(x^j, y^j) \sim (z^j, w^j)$ for all $j < m] \Rightarrow$ not $(x^m, y^m) \prec (z^m, w^m)$, and $(x, y) \sim (y, x)$ for all $x, y \in X$.

6.5. $x - y \sim^* z - w \Rightarrow$ not $x - y \prec^* z - w \Rightarrow$ not $x - z \prec^* y - w$ (by (6.2)) $\Rightarrow$ $[x - z \sim^* y - w$ or $y - w \prec^* x - z]$ the latter of which gives $z - x \prec^* w - y$

by (6.1). Also, $x - y \sim^* z - w \Rightarrow$ not $z - w \prec^* x - y \Rightarrow$ not $z - x \prec^* w - y$ by (6.2). Therefore $x - y \sim^* z - w \Rightarrow x - z \sim^* y - w$. (And so forth.)

6.9. For negative transitivity, not $x \prec y \Rightarrow y - x \preccurlyeq^* x - x$ and not $y \prec z \Rightarrow z - y \preccurlyeq^* y - y$. Hence $z - y \preccurlyeq^* x - x$, then $z - x \preccurlyeq^* y - x$, then $z - x \preccurlyeq^* x - x$.

6.15a. Asymmetry of $\prec^*$ is immediate from asymmetry of $\prec$. For negative transitivity, not $x - y \prec^* z - w \Rightarrow$ not $f(x, w) \prec f(z, y) \Rightarrow f(z, y) \preccurlyeq f(x, w)$, and not $z - w \prec^* s - t \Rightarrow$ not $f(z, t) \prec f(s, w) \Rightarrow f(s, w) \preccurlyeq f(z, t)$. Using $C3$, $C2$, $C3$, $C2$, and $C3$ again, $f(f(s, y), f(w, z)) \sim f(f(s, w), f(y, z)) \preccurlyeq f(f(z, t), f(y, z)) \sim f(f(z, y), f(t, z)) \preccurlyeq f(f(x, w), f(t, z)) \sim f(f(x, t), f(w, z))$: by transitivity, $f(f(s, y), f(w, z)) \preccurlyeq f(f(x, t), f(w, z))$, whence $f(s, y) \preccurlyeq f(x, t)$ by $C2$, which says that $s - t \preccurlyeq^* x - y$, or not $x - y \prec^* s - t$.

6.16. To show that $f(f(x, y), f(z, w)) \sim f(f(x, z), f(y, w))$ let $a = f(x, y)$, $b = f(z, w)$, $c = f(x, z)$, $d = f(y, w)$. We are to show that $a - d \sim^* c - b$. The permutation condition of B_6 holds for $x - a \sim^* a - y$, $w - b \sim^* b - z$, $c - z \sim^* x - c$, $d - y \sim^* w - d$, $b - d \mathbin{?} c - a$, $a - d \mathbin{?^*} c - b$. If $? = \prec^*$ then not $a - d \prec^* c - b$ by B_6, and hence $c - b \prec^* a - d$ by B_6, and hence $c - a \prec^* b - d$ by (6.2), which contradicts $b - d \prec^* c - a$. Similarly $c - a \prec^* b - d$ can't hold. Hence $b - d \sim^* c - a$, which by B_6 yields $a - d \sim^* c - b$.

7.1. Since $(x_1, \ldots, x_n) \sim (x_2, \ldots, x_n, x_1) \sim \cdots \sim (x_n, x_1, \ldots, x_{n-1})$, $\sum u_i(x_i) = \sum \rho(x_i)/n$.

7.4b. $x \preccurlyeq (x_1, \ldots, x_{n-1}, y_n) \preccurlyeq (x_1, \ldots, x_{n-2}, y_{n-1}, y_n) \preccurlyeq \cdots \preccurlyeq y$.

7.5. For the last part: $A = \{a, b\}$, $n = 2$, $(a, a) \prec (b, a)$, $(a, a) \prec (a, b)$, $(a, b) \prec (b, b)$, $(b, a) \prec (b, b)$ and $(a, a) \sim (b, b)$.

7.13k. Suppose $\alpha \neq 0$, $\alpha = M/N$ where M, N are nonzero integers. Then $x \sim y \Leftrightarrow Mx \sim My$ by e and j. Since $\alpha N = M$, $x \sim y \Leftrightarrow N\alpha x \sim N\alpha y$, and by e and j, $N\alpha x \sim N\alpha y \Leftrightarrow \alpha x \sim \alpha y$.

8.1. Expected net profit maximized at about $x = 235000$.

8.4a. $v(z) = 3$, $v(w) = 9$. (d) $v = 5u + 5$.

8.5. $\alpha = .4$.

8.6d. Show that (conditions 1, 2, not 3) $\Rightarrow$ not 4. In violation of 3 assume that $P \prec Q \prec R$ and $Q \preccurlyeq \alpha P + (1 - \alpha)R$ for all $\alpha \in (0, 1)$. Show first that, for every $\alpha, \beta \in (0, 1)$, $(1 - \beta)Q + \beta R \preccurlyeq (1 - \alpha)(1 - \beta)P + [\beta + (1 - \beta)\alpha]R \prec (1 - \alpha)(1 - \beta)Q + [\beta + (1 - \beta)\alpha]R$. (Note that $Q \prec \alpha P + (1 - \alpha)R$ for all $\alpha \in (0, 1)$. Why?) Suppose that $S \prec T \Leftrightarrow u(S) < u(T)$ for all $S, T \in \mathcal{P}_s$. Let $f(\beta) = u((1 - \beta)Q + \beta R)$, $g(\beta) = u((1 - \beta)P + \beta R)$ for all $\beta \in (0, 1)$ and note that if $\beta < \gamma$ then $f(\beta) \leq g(\gamma) < f(\gamma)$, with $\gamma = \beta + (1 - \beta)\alpha$. Show that $\sup\{f(\beta): \beta < \gamma\} < f(\gamma)$ so that f is discontinuous at every point in (0, 1). This is impossible (why?) and therefore such a u does not exist. Then use Theorem 3.1 to show that condition 4 is false.

8.7. See Exercise 6c.

8.11a. For (A, C), $P(\$30) = .27$, $P(\$70) = .63$, $P(\$80) = .03$, $P(\$120) = .07$. The theory of this chapter does not say that (B, D) will be preferred.

8.13. No. Yes.

8.16a. $y \doteq \$15000$.

8.16b. $y' = y$. Given A, he would sell it for an amount with the same utility.

8.16c. \$0 ~ (\$40000 − z with pr. 1/2 or \$0 − z with pr. 1/2). $z \doteq$ \$18000. If he paid \$18000 for A he would be taking a 50-50 gamble between net increments of \$22000 and −\$18000, which has a utility of about zero, which is what he started with.

8.16d. Of course not. In the two situations he is considering different amounts of total wealth. He would sell it for \$18000 or more.

8.16e. (\$25000 with pr. 1/2 or - \$15000 with pr. 1/2) ~ w − \$15000. $w \doteq$ \$20000.

8.16f. With y as in part a, y ~ (\$0 − r with pr. 1/4 or \$40000 − r with pr. 1/2 or \$80000 − r with pr. 1/4).

8.16g. (\$25000 with pr. 1/2 or −\$15000 with pr. 1/2) ~ (\$0 − \$15000 − s with pr. 1/4 or \$40000 − \$15000 − s with pr. 1/2 or \$80000 − \$15000 − s with pr. 1/4). $s \doteq$ \$12500.

9.1. For the converse of $B3$ suppose $\alpha P + (1 - \alpha)R \approx \alpha Q + (1 - \alpha)R$ with $\alpha \in (0, 1)$ and not $P \approx Q$. Then, for example, $P \sim T$ and $T \prec Q$. By $B2$, $\alpha P + (1 - \alpha)R \sim \alpha T + (1 - \alpha)R$. By $B1$, $\alpha T + (1 - \alpha)R \prec \alpha Q + (1 - \alpha)R$. These contradict $\alpha P + (1 - \alpha)R \approx \alpha Q + (1 - \alpha)R$.

9.3. Suppose $P \sim Q$, $Q \sim R$. Then $\frac{1}{2}P + \frac{1}{2}R \sim \frac{1}{2}Q + \frac{1}{2}R$ and $Q \sim \frac{1}{2}Q + \frac{1}{2}R$ by $B2$. Therefore $\frac{1}{2}(\frac{1}{2}P + \frac{1}{2}R) + \frac{1}{2}Q \sim \frac{1}{2}R + \frac{1}{2}Q$ by $B4$. Therefore $\frac{1}{2}P + \frac{1}{2}R \sim R$ by $(B1, B2)$. Hence $P \sim R$ by $(B1, B2)$.

9.5. If $P \sim Q$ but $P \not\approx Q$ then by a \$1 change in a consequence of P or else a small change in two probabilities in P it would seem possible to get a P^* that is indifferent to Q but either preferred to P or less preferred than P.

9.7. Suppose $(z - y)^2 > \inf\{(x - y)^2 : x \in \bar{X}\}$ for all $z \in \bar{X}$. Then there is a sequence $z_1, z_2, \ldots$ in $\bar{X}$ such that $(z_1 - y)^2 > (z_2 - y)^2 > \cdots$ and $\inf\{(z_n - y)^2 : n = 1, 2, \ldots\} = \inf\{(x - y)^2 : x \in \bar{X}\}$. Then there is a z such that every open set in Re^n that contains z must contain some z_n. Since the closure of $\bar{X}$ is $\bar{X}$ it follows that $z \in \bar{X}$ and that $(z - y)^2 = \inf\{(z_n - y)^2 : n = 1, 2, \ldots\}$, which contradict the original supposition.

9.9d. $\{(\alpha, \alpha) : 0 \leq \alpha \leq 1\}$.

9.13(2). $\lambda_i \geq 0$ for all i and $\sum \lambda_i > 0$.

10.4c. (0, 1) except for 1/2, 1/3, 1/4, ...

10.5. All subsets of Re that contain either a countable number of elements or all but a countable number of elements in Re.

10.7. If $\sup\{r + s : r \in R, s \in S\} < \sup R + \sup S$ then $\sup\{r + s : \cdots\} < r + s$ for some $r \in R$ and $s \in S$, so that $r + s < r + s$.

10.9. If $\sup\{\sum_1^\infty \alpha_j\beta_{ij} : i = 1, 2, \ldots\} < \sum_1^\infty \alpha_j[\sup\{\beta_{ij} : i = 1, 2, \ldots\}]$ then $\sum_{j=1}^\infty \alpha\,\beta_{ij} + \epsilon < \sum_{j=1}^m \alpha_j[\sup\{\beta_{ij} : i = 1, 2, \ldots\}]$ for some m, some $\epsilon > 0$, and $t = 1, 2, \ldots$. Let β_{i_jj} be such that $\sup\{\beta_{ij} : i = 1, 2, \ldots\} \leq \beta_{i_jj} + \epsilon/M$ for $j = 1, 2, \ldots, m$. Let s be the largest such i_j. Then $\sup\{\beta_{ij} : i = 1, 2, \ldots\} \leq \beta_{sj} + \epsilon/M$ for $j = 1, 2, \ldots, m$, and $\sum_{j=1}^m \alpha_j \sup\{\beta_{ij} : i = 1, 2, \ldots\} \leq \sum_{j=1}^m \alpha_j[\beta_{sj} + \epsilon/M] < \sum_{j=1}^m \alpha_j\beta_{sj} + \epsilon$. Hence $\sum_{j=1}^\infty \alpha_j\beta_{tj} < \sum_{j=1}^m \alpha_j\beta_{sj}$ for all t and hence for $t = s$, which is impossible.

Suppose $\sum_{j=1}^\infty \alpha_j \sup\{\beta_{ij} : i = 1, 2, \ldots\} < \sup\{\sum_{j=1}^\infty \alpha_j\beta_{ij} : i = 1, 2, \ldots\}$. Then $\sum_{j=1}^\infty \alpha_j \sup\{\beta_{ij} : i = 1, 2, \ldots\} + \epsilon < \sum_{j=1}^\infty \alpha_j\beta_{kj}$ for some k, some $\epsilon > 0$ and all n. Hence $\sum_{j=1}^n \alpha_j \sup\{\beta_{ij} : i = 1, 2, \ldots\} < \sum_{j=1}^m \alpha_j\beta_{kj}$ for some m and all n. But $\sum_{j=1}^m \alpha_j\beta_{kj} \leq \sum_{j=1}^m \alpha_j \sup\{\beta_{ij} : i = 1, 2, \ldots\}$ and a contradiction is obtained.

10.13. Let $A_1, A_2, \ldots$ be mutually disjoint elements in $\mathcal{A}$ with $\bigcup_{j=1}^{\infty} A_j \in \mathcal{A}$. Then $\sum_{i=1}^{\infty} \alpha_i P_i(\bigcup_{j=1}^{\infty} A_j) = \sum_i \alpha_i \sum_j P_i(A_j) = \sum_i \alpha_i \sup\{\sum_{j=1}^{n} P_i(A_j): n = 1, 2, \ldots\} = \sup\{\sum_{i=1} \alpha_i(\sum_{j=1}^{n} P_i(A_j)): n = 1, 2, \ldots\}$ [by Exercise 9] $= \sup\{\sum_{j=1}^{n} \sum_{i=1}^{\infty} \alpha_i P_i(A_j): n = 1, 2, \ldots\} = \sum_j \sum_i \alpha_i P_i(A_j)$.

10.15. Let $f_1, f_2, \ldots$ and $g_1, g_2, \ldots$ be sequences of simple $\mathcal{A}$-measurable functions satisfying (1) and (2) of Definition 10.11, and suppose that $\sup\{E(f_n, P)\} < \sup\{E(g_n, P)\}$. Then $E(f_n, P) + \epsilon < E(g_m, P)$ for some m, some $\epsilon > 0$, and all n. Let $A_n = \{x: g_m(x) \leq f_n(x) + \epsilon/2\}$ so that $A_1 \subseteq A_2 \subseteq \cdots$ and $P(A_1) \leq P(A_2) \leq \cdots$. Also, $X = \bigcup_{n=1}^{\infty} A_n$ so that $1 = P(\bigcup_{n=1}^{\infty} A_n) = \sup\{P(A_n): n = 1, 2, \ldots\}$ by Lemma 10.2. Let $M = \sup\{g_m(x) - f_m(x): x \in X\}$. Then for $n \geq m$, $E(g_m, P) - E(f_n, P) = E(g_m - f_n, P) \leq M[1 - P(A_n)] + P(A_n)\epsilon/2$, the equality coming from Exercise 17. As n gets large, the right side of this approaches $\epsilon/2$ and hence $E(g_m, P) - E(f_n, P) < \epsilon$ for some n, a contradiction.

10.18. $E(f, \alpha P + (1 - \alpha)Q) = \sup\{E(f_n, \alpha P + (1 - \alpha)Q): n = 1, 2, \ldots\} = \sup\{\alpha E(f_n, P) + (1 - \alpha)E(f_n, Q)\} = \alpha \sup\{E(f_n, P)\} + (1 - \alpha)\sup\{E(f_n, Q)\}$ [Exercises 6, 8, 22*a*] $= \alpha E(f, P) + (1 - \alpha)E(f, Q)$.

10.21c. Use the results of Exercises 21*a* and 19.

10.22a. Let $a < f(x) \leq b$, $a < g(x) \leq b$ for all x. Let $A_{i,n}$ and f_n be defined by (10.9) and (10.10), and let $B_{i,n} = \{x: a + (i - 1)(b - a)/n < g(x) \leq a + i(b - a)/n\}$ and $g_n(x) = a + (i - 1)(b - a)/n$ for all $x \in B_{i,n}$. Let $c_{in} = a + (i - 1)(b - a)/n$. $E(f_n, P) = \sum_1^n P(A_{i,n})c_{in}$ and $E(g_n, P) = \sum_1^n P(B_{i,n})c_{in}$. $E(f_n, P) \leq E(g_n, P)$ follows from $\sum_1^k P(B_{i,n}) \leq \sum_1^k P(A_{i,n})$ for $k = 1, \ldots, n$, which in turn follows from $P(f(x) \leq g(x)) = 1$.

10.25. $\{x: x \in X, y \prec x \prec z\} = [\{x: y \prec x\}^c \cup \{x: x \prec z\}^c]^c$.

10.26. For Theorem 10.2 let $P(x) = 0$ for all x, $R = \frac{1}{2}P + \frac{1}{2}1$.

10.27. Show that if $\mathcal{G}^* \subseteq \mathcal{G}$ has elements weakly ordered by $\subset$ then $\mathcal{S}^* = \bigcup_{\mathcal{G}^*} \mathcal{S}$ is in $\mathcal{G}$.

11.2c. It can be true for some pair $P, Q \in \mathcal{T}$ that $P \sim Q$ when $P_i \prec_i Q_i$ and $P_i^c = Q_i^c$. For some other P, Q pair, $P \prec Q$.

11.2d. Let R^k, $k = 1, 2, \ldots, m - 1$, be such that $R_1^1 = Q_1$, $R_1^{1c} = P_1^c$; $R_k^k = Q_k$, $R_k^{kc} = R_k^{k-1c}$ for $k = 2, 3, \ldots, m - 1$. Then $P \preccurlyeq R^1, R^1 \preccurlyeq R^2, \ldots, R^{m-1} \preccurlyeq Q$.

11.2e. Let $n = 2$, $(x_1, x_2) \sim (y_1, x_2) \sim (x_1, y_2) \prec (y_1, y_2) \sim P$ for all P on $\{x_1, y_1\} \times \{x_2, y_2\}$ that are not one-point measures. Show that $\preccurlyeq_1$ and $\preccurlyeq_2$ are transitive and connected, that $x_1 \prec_1 y_1$, $x_2 \sim_2 y_2$, but $(x_1, x_2) \sim (y_1, x_2)$.

11.2j. An example where A and D hold but $\preccurlyeq_1$ is not connected: $X = \{x_1, y_1\} \times \{x_2, y_2\}$, $(x_1, x_2) \prec (y_1, x_2) \prec (x_1, y_2) \prec (y_1, y_2) \prec P \sim Q$ for any P, Q that are not one-point measures.

11.2k. With $a \prec_i b$ there are $R, S \in \mathcal{T}$ such that $R \prec S$, $R_i = a$, $S_i = b$, and $R_i^c = S_i^c$. Let $T = \frac{1}{2}R + \frac{1}{2}Q$, $T' = \frac{1}{2}S + \frac{1}{2}P$. $T \sim T'$ by D. With $R \prec S$, if $Q \sim P$, then $\frac{1}{2}R + \frac{1}{2}Q \prec \frac{1}{2}S + \frac{1}{2}Q \sim \frac{1}{2}S + \frac{1}{2}P$, or $T \prec T'$, a contradiction. $P \preccurlyeq Q$ by definition of $\prec_i$. Hence $P \prec Q$ since $P \sim Q$ is false.

11.2m. Let $X = \{x_1, y_1\} \times \{x_2, y_2\}$, take $P \preccurlyeq Q \Leftrightarrow (P_1(x_1), P_2(x_2)) \preccurlyeq (Q_1(x_1), Q_2(x_2))$ and on $[0, 1] \times [0, 1]$ take $(\alpha, \beta) \prec (\gamma, \delta)$ if and only if $\alpha < \gamma$ or $[\alpha = \gamma$ and $\alpha\beta < \gamma\delta]$, with $\preccurlyeq$ on $[0, 1]^2$ transitive and connected. D holds since $(\alpha, \beta) \sim (\alpha, \beta)$. Moreover, $(0, 0) \sim (0, 1)$ and $(1, 0) \prec (1, 1)$: that is, $(y_1, y_2) \sim (y_1, x_2)$ and $(x_1, y_2) \prec (x_1, x_2)$, from which it follows that $y_2 \prec_2 x_2$ but $(y_1, y_2) \sim (y_1, x_2)$. B says

that if $(\alpha, \beta) \prec (\gamma, \delta)$ and $(j, k) \in [0, 1]^2$ and $t \in (0, 1)$, then $(t\alpha + (1-t)j,$ $t\beta + (1-t)k) \prec (t\gamma + (1-t)j, t\delta + (1-t)k)$, which is easily seen to be true.

11.2o. Let $P_i^c = Q_i^c = R_i^c$, $P_i \preccurlyeq_i Q_i$, $\alpha \in (0, 1)$. Then $P \preccurlyeq Q$. By B and C, $\alpha P + (1-\alpha)R \preccurlyeq \alpha Q + (1-\alpha)R$. Since $\alpha P_i^c + (1-\alpha)R_i^c = \alpha Q_i^c + (1-\alpha)R_i^c$, if $\alpha Q_i + (1-\alpha)R_i \prec_i \alpha P_i + (1-\alpha)R_i$ then, by Exercises 2*f* and 2*l*, $\alpha Q + (1-\alpha)R \prec \alpha P + (1-\alpha)R$, a contradiction. Since $\preccurlyeq_i$ is connected (Exercise 2*n*), $\alpha P_i + (1-\alpha)R_i \preccurlyeq_i \alpha Q_i + (1-\alpha)R_i$. For the latter part of the theorem take $P_i \prec_i Q_i$.

11.3a. $E(f, P) = \sum_X f_i(x_i)P(x_1, \ldots, x_n) = \sum_{X_i} f_i(x_i)P(X_1 \times \cdots \times X_{i-1} \times \{x_i\} \times X_{i+1} \times \cdots \times X_n) = \sum_{X_i} f_i(x_i)P_i(x_i) = E(f_i, P_i)$.

11.9b. The condition of part *a* leads directly to $u(x_1, \ldots, x_i, x_{i+1}^0, \ldots, x_n^0) + u(x_1^0, \ldots, x_{i-1}^0, x_i, x_{i+1}, x_{i+2}^0, \ldots, x_n^0) = u(x_1, \ldots, x_{i+1}, x_{i+2}^0, \ldots, x_n^0) + u(x_1^0, \ldots, x_{i-1}^0, x_i, x_{i+1}^0, \ldots, x_n^0)$. Let $u_i(x_i, x_{i+1}) = u(x_1^0, \ldots, x_{i-1}^0, x_i, x_{i+1}, x_{i+2}^0, \ldots, x_n^0) - u(x_1^0, \ldots, x_i^0, x_{i+1}, x_{i+2}^0, \ldots, x_n^0)$ for $i = 1, \ldots, n-2$, and $u_{n-1}(x_{n-1}, x_n) = u(x_1^0, \ldots, x_{n-2}^0, x_{n-1}, x_n)$.

12.1. $P'(\{s : s(f) \in A\} \cap \{s : s(f) \in A'\})/P_g(A')$.

12.6b. $[v(f, s_2) - v(g, s_2)]a = P^*(s_2)[u(\text{win}) - u(\text{lose})]$, $[v(g, s_3) - v(f, s_3)]a = P^*(s_3)[u(\text{win}) - u(\text{lose})]$, and so forth.

12.9. He would rather marry Alice but should propose to Betsy. Use $a = P^*(s_1)[4 - 3]$, $2a = P^*(s_2)[4 - 0]$, and $4a = P^*(s_3)[3 - 0]$.

13.4. The simplest example is $\mathcal{B}^2 = \{A, B\}$, $\mathcal{B}^3 = \{A_1, A_2, B\}$, and $\mathcal{B}^4 = \{A_1, A_2, B_1, B_2\}$. If A is selected from $\mathcal{B}^2$ then B must be selected from $\mathcal{B}^3$, but $A \cup B = S$. If B is selected from $\mathcal{B}^2$ and A_1 (or A_2) is selected from $\mathcal{B}^3$ then A_2 (or A_1) must be selected from $\mathcal{B}^4$. But $B \cup A_1 \cup A_2 = S$.

13.5. Let $B = \bigcap_{\mathcal{A}} A$ with $B \neq \varnothing$. If $P^*(B) = 0$, then $P^*(B^c) = 1$ so that $B^c \in \mathcal{A}$, which contradicts $B = \bigcap_{\mathcal{A}} A$. Hence $P^*(B) = 1$. If B has more than one element then B can be partitioned into C and D with $P(C) = 0$ and $P(D) = 1$, $D \subset B$, which contradicts $B = \bigcap_{\mathcal{A}} A$.

13.10. Let u be such that $u(x_*) = \inf\{u(x) : x \in X\} = 0$ and $u(x^*) = \sup\{u(x) : x \in X\} = 1$. Given $\mathbf{P} \in \mathcal{H}$ let $A_{1,n} = \{s : 0 \le E(u, \mathbf{P}(s)) \le 1/n\}$ and $A_{i,n} = \{s : (i-1)/n < E(u, \mathbf{P}(s)) \le i/n\}$ for $i = 2, \ldots, n$. Define $\mathbf{P}_n$ and $\mathbf{Q}_n$ in $\mathcal{H}$ by $\mathbf{P}_n(s) = [(i-1)/n]x^* + [(n-i+1)/n]x_*$ for all $s \in A_{i,n}$ and $\mathbf{Q}_n(s) = (i/n)x^* + [(n-i)/n]x_*$ for all $s \in A_{i,n}$; $i = 1, \ldots, n$. It follows from *B*7 that $\mathbf{P}_n \preccurlyeq \mathbf{P} \preccurlyeq \mathbf{Q}_n$ and hence that $v(\mathbf{P}_n) \le v(\mathbf{P}) \le v(\mathbf{Q}_n)$ for all n, where v is as defined in the proof of *S*2. Hence, by (13.9) for all horse lotteries in $\mathcal{H}_0$, $E[E(u, \mathbf{P}_n(s)), P^*] \le v(\mathbf{P}) \le E[E(u, \mathbf{Q}_n(s)), P^*]$ for all n.

13.12. Let $\mathbf{Q}$ be as defined following (13.12). Assume $c = 0$, $d = 1$ for convenience. [If $c = d$ the result is immediate.] Let $\mathbf{R}_i = R_i$ on S with $E(u, R_i) = i/4$ for $i = 1, 2, 3$. Since $0 \le E(u, \mathbf{Q}(s)) \le 1$ for all s, $1/4 \le E(u, \frac{1}{2}\mathbf{Q}(s) + \frac{1}{2}R_2) \le 3/4$ for all $s \in S$. Therefore $R_1 \preccurlyeq \frac{1}{2}\mathbf{Q}(s) + \frac{1}{2}R_2 \preccurlyeq R_3$ for all $s \in S$. Hence, by *B*7, $\mathbf{R}_1 \preccurlyeq \frac{1}{2}\mathbf{Q} + \frac{1}{2}\mathbf{R}_2 \preccurlyeq \mathbf{R}_3$. By (13.10) and (13.11), $v(\mathbf{R}_1) \le \frac{1}{2}v(\mathbf{R}_2) + \frac{1}{2}v(\mathbf{Q}) \le v(\mathbf{R}_3)$. Then by (13.9) for $\mathcal{H}_0$, $1/4 \le \frac{1}{2}v(\mathbf{Q}) + 1/4 \le 3/4$, or $0 \le v(\mathbf{Q}) \le 1$.

13.15a. Given $\epsilon > 0$ let $B(\epsilon) = \{s : \alpha E(u, \mathbf{P}(s)) + (1-\alpha)E(u, \mathbf{R}(s)) \ge 1 - \epsilon\}$ and $C(\epsilon) = \{s : s \in B(\epsilon)$, and $E(u, \mathbf{P}(s)) < 1 - \epsilon$ or $E(u, \mathbf{R}(s)) < 1 - \epsilon\}$. Let $\delta = \alpha(1-\alpha)\epsilon$. Then, if $s \in C(\epsilon)$, s cannot be in $B(\delta)$ since $\alpha(1-\epsilon) + (1-\alpha) < 1 - \epsilon$ and $\alpha + (1-\alpha)(1-\epsilon) < 1 - \epsilon$. Hence the only elements in $B(\epsilon)$

for any $\epsilon > 0$ that can contribute to $\inf\{P^*\{\alpha E(u, \mathbf{P}(s)) + (1 - \alpha)E(u, \mathbf{R}(s)) \geq 1 - \epsilon\}: \epsilon > 0\}$ are those for which both $E(u, \mathbf{P}(s)) \geq 1 - \epsilon$ and $E(u, \mathbf{R}(s)) \geq 1 - \epsilon$. As a consequence, $\inf\{P^*\{\alpha E(u, \mathbf{P}(s)) + (1 - \alpha)E(u, \mathbf{R}(s)) \geq 1 - \epsilon\}: \epsilon > 0\} = \inf\{P^*(\{E(u, \mathbf{P}(s)) \geq 1 - \epsilon\} \cap \{E(u, \mathbf{R}(s)) \geq 1 - \epsilon\}): \epsilon > 0\}$.

13.15c. Let $\mathbf{P}, \mathbf{Q}, \mathbf{R} \in \mathcal{H}$ be as follows. On the even integers, $E(u, \mathbf{P}(s)) = s/(1 + s)$, $E(u, \mathbf{Q}(s)) = E(u, \mathbf{R}(s)) = 0$. On the odd integers, $\mathbf{P}(s)(\frac{1}{2}) = 1$, and $E(u, \mathbf{Q}(s)) = E(u, \mathbf{R}(s)) = s/(1 + s)$. Then $v(\mathbf{P}) = [\frac{1}{2} + (\frac{1}{2})(\frac{1}{2})] + \frac{1}{2} = 5/4$, $v(\mathbf{Q}) = \frac{1}{2} + \frac{1}{2} = 1$, $v(\frac{1}{2}\mathbf{P} + \frac{1}{2}\mathbf{R}) = \frac{1}{2}(3/4) + \frac{1}{2}(\frac{1}{2}) + 0 = 5/8$, and $v(\frac{1}{2}\mathbf{Q} + \frac{1}{2}\mathbf{R}) = \frac{1}{2}(\frac{1}{2}) + \frac{1}{2}(\frac{1}{2}) + \frac{1}{2} = 1$. Hence $\mathbf{Q} \prec \mathbf{P}$ and $\frac{1}{2}\mathbf{P} + \frac{1}{2}\mathbf{R} \prec \frac{1}{2}\mathbf{Q} + \frac{1}{2}\mathbf{R}$.

14.2b. Consider $pr \prec^* s$, $ps \prec^* qr$, $rs \prec^* pt$, and $qt \prec^* prs$.

14.4. $B \subseteq C \Rightarrow C = B \cup (C/B)$. By $(F1, F3)$, $\varnothing \preccurlyeq^* C/B$ so that, by $(F3, F4)$, $\varnothing \cup B \preccurlyeq^* (C/B) \cup B$, or $B \prec^* C$. If $S \prec^* C$ then $C \cup C^c \prec^* C$, which by $F4$ implies $C^c \prec^* \varnothing$, contradicting $F1$.

14.5e. $A \prec^* C \Rightarrow B \prec^* D \Rightarrow A \cup B \prec^* C \cup D$ by $C3(\prec^*)$, a contradiction.

14.10. Let $F5$ hold. Then, by $F5$ directly, $\prec^*$ is fine. If $A \prec^* B$, it follows easily from $F5$ and the other properties of $\prec^*$ that $A \cup D \prec^* B$ for some D for which $A \cap D = \varnothing$ and $\varnothing \prec^* D$. A similar result is obtained if $B \prec^* A$. Hence, if the "$A \preccurlyeq^* B \cup C$ and $B \preccurlyeq^* A \cup D$ for all $\cdots$" conditions of tightness hold, then neither $A \prec^* B$ nor $B \prec^* A$, which requires $A \sim^* B$ by $F3$.

On the other hand, suppose $F1$–$F4$ hold and $\prec^*$ is fine and tight. Take $A \prec^* B$. Suppose, for all $B_1 \subseteq B$ for which $B_1 \prec^* B$, $B_1 \preccurlyeq^* A$. Then consider $\varnothing \prec^* D$ and $A \cap D = \varnothing$. By fineness, it follows that there is a $B_2 \subseteq B$ such that $\varnothing \prec^* B_2 \preccurlyeq^* D$. Then since $\varnothing \prec^* B_2$, $B/B_2 \prec^* B$ so that $B/B_2 \preccurlyeq^* A$, which along with $B_2 \preccurlyeq^* D$ gives $B \preccurlyeq^* A \cup D$ by $C3$. Tightness then requires that $A \sim^* B$, which is false. Hence, with $A \prec^* B$, there is a $B_1 \subseteq B$ for which $A \prec^* B_1 \prec^* B$. Since $\varnothing \prec^* B/B_1$ and $\prec^*$ is fine, there is a partition $\{C_1, \ldots, C_m\}$ of S with $C_i \preccurlyeq^* B/B_1$ for each i. Along with $A \prec^* B_1$, this gives $A \cup C_i \prec^* B$ by $C3(\prec^*)$.

14.11c. Given A, B let the "whenever" conditions of tightness hold. If no C satisfies $B \cap C = \varnothing$ and $\varnothing \prec^* C$ then $B \sim^* S$ so that $A \preccurlyeq^* B$. If some C satisfies $\varnothing \prec^* C$ and $B \cap C = \varnothing$ then, for any such C, either $P_1(A_1) < P_1(B_1) + P_1(C_1)$ or $P_1(A_1) = P_1(B_1) + P_1(C_1)$ and $P_2(A_2) \leq P_2(B_2) + P_2(C_2)$. If $P_2(C_2) = 0$ for all such C_2 then $P_2(B_2) = 1$ which insures $P_2(A_2) \leq P_2(B_2)$, and since $P_1(C_1) > 0$ can be made arbitrarily small, we get also $P_1(A_1) \leq P_1(B_1)$. Hence, if $P_2(C_2) \equiv 0$, $A \preccurlyeq^* B$. If $P_2(C_2) > 0$ for some such C_2, then we can take a C_1 with $P_1(C_1) = 0$ and get $P_1(A_1) \leq P_1(B_1)$, where, if equality holds, it must then be true that $P_2(A_2) \leq P_2(B_2)$. Again, $A \preccurlyeq^* B$. By a similar proof we get $B \preccurlyeq^* A$ for all cases. Hence $A \sim^* B$.

14.13c. Let $A = [0, \frac{1}{2}) \cup S_2$, $B = (\frac{1}{2}, 1] \cup S_2$. If $\prec^*$ is tight, then $A \sim^* B$. But $B \prec^* A$ by the definition of $\prec^*$.

14.17. If A is not null then, with $x \prec y$ and $f = x$ on A, $f = y$ on A^c and $g = y$ on S, $f \prec g$ given A by $P3$ so that $f \prec g$ by Lemma 14.1. But if $A \sim^* \varnothing$ then $f \sim g$. Therefore $\varnothing \sim^* A$ implies that A is null.

14.20b. For $P6$, suppose $f \prec g$ or $w(f) < w(g)$. Let $w(g) - w(f) = d$. Take a partition $\{\{1, n, 2n, \ldots\}, \{2, n + 1, 2n + 1, \ldots\}, \ldots, \{n - 1, 2n - 1, \ldots\}\}$ with n large enough so that $2 < dn$, and use the answer to part (*a*).

REFERENCES

The references given here are those cited in the text. For a more complete bibliography see Fishburn (1968) and the other surveys cited in its introduction.

Adams, E. W. (1965). Elements of a theory of inexact measurement. *Philosophy of Science*, **32**, 205–228.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'ecole Americaine. *Econometrica*, **21**, 503–546.

Anscombe, F. J. and Aumann, R. J. (1963). A definition of subjective probability. *Annals of Mathematical Statistics*, **34**, 199–205.

Armstrong, W. E. (1939). The determinateness of the utility function. *Economic Journal*, **49**, 453–467.

——— (1950). A note on the theory of consumer's behaviour. *Oxford Economic Papers*, **2**, 119–122.

Arrow, K. J. (1959). Rational choice functions and orderings. *Economica*, **26**, 121–127.

——— (1966). Exposition of the theory of choice under uncertainty. *Synthese*, **16**, 253–269.

Aumann, R. J. (1962). Utility theory without the completeness axiom. *Econometrica*, **30**, 445–462.

——— (1964). Utility theory without the completeness axiom: a correction. *Econometrica*, **32**, 210–212.

——— (1964). Subjective programming. In M. W. Shelly and G. L. Bryan (Eds.), *Human judgments and optimality*. Wiley, New York.

Bernoulli, D. (1738). Specimen theoriae novae de mensura sortis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, **5**, 175–192. Translated by L. Sommer, *Econometrica*, **22** (1954), 23–36.

Birkhoff, G. (1948). *Lattice theory*, Rev. Ed. American Mathematical Society, New York.

Blackwell, D. and Girshick, M. A. (1954). *Theory of games and statistical decisions*. Wiley, New York.

Blaschke, W. (1928). Topologische Fragen der Differentialgeometrie. I. Thomsens Sechseckgewebe. Zueinander diagonale Netze. *Mathematische Zeitschrift*, **28**, 150–157.

Chernoff, H. (1954). Rational selection of decision functions. *Econometrica*, **22**, 422–443.

Chipman, J. S. (1960). The foundations of utility. *Econometrica*, **28**, 193–224.

Coombs, C. H. (1964). *A theory of data.* Wiley, New York.

Cramer, H. (1956). A theorem on ordered sets of probability distributions. *Theory of Probability and Its Applications*, **1**, 16–21.

Debreu, G. (1954). Representation of a preference ordering by a numerical function. In R. M. Thrall, C. H. Coombs, and R. L. Davis (Eds.), *Decision processes.* Wiley, New York.

——— (1959). *Theory of value.* Wiley, New York.

——— (1960). Topological methods in cardinal utility theory. In K. J. Arrow, S. Karlin, and P. Suppes (Eds.), *Mathematical methods in the social sciences, 1959.* Stanford University Press, Stanford, California.

——— (1964). Continuity properties of Paretian utility. *International Economic Review*, **5**, 285–293.

de Finetti, B. (1937). La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré*, **7**, 1–68. Translated by H. E. Kyburg in Kyburg and Smokler (1964).

Diamond, P. A. (1965). The evaluation of infinite utility streams. *Econometrica*, **33**, 170–177.

Dubins, L. E. and Savage, L. J. (1965). *How to gamble if you must: inequalities for stochastic processes.* McGraw-Hill, New York.

Eilenberg, S. (1941). Ordered topological spaces. *American Journal of Mathematics*, **63**, 39–45.

Ellsberg, D. (1954). Classic and current notions of "measurable utility." *Economic Journal*, **64**, 528–556.

——— (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, **75**, 643–669.

Fishburn, P. C. (1964). *Decision and value theory.* Wiley, New York.

——— (1967). Methods of estimating additive utilities. *Management Science*, **13**, 435–453.

——— (1968). Utility theory. *Management Science*, **14**, 335–378.

——— (1969). A general theory of subjective probabilities and expected utilities. *Annals of Mathematical Statistics*, **40**, 1419–1429.

Friedman, M. and Savage, L. J. (1948). The utility analysis of choices involving risk. *Journal of Political Economy*, **56**, 279–304.

——— and ——— (1952). The expected-utility hypothesis and the measurability of utility. *Journal of Political Economy*, **60**, 463–474.

Frisch, R. (1926). Sur une probleme d'économie pure. *Norsk Mathematish Forenings Skrifter*, **1**, 1–40.

——— (1964). Dynamic utility. *Econometrica*, **32**, 418–424.

Fuchs, L. (1963). *Partially ordered algebraic systems.* Addison-Wesley, Reading, Massachusetts. Copyright: Akadémia Kiadó, Budapest.

Galanter, E. (1962). The direct measurement of utility and subjective probability. *American Journal of Psychology*, **75**, 208–220.

Goldman, A. J. (1956). Resolution and separation theorems for polyhedral convex sets. In H. W. Kuhn and A. W. Tucker (Eds.), *Linear inequalities and related systems, Annals of Mathematics Study 38.* Princeton University Press, Princeton, New Jersey.

Halmos, P. R. (1950). *Measure theory*. Van Nostrand, New York.

Hausner, M. (1954). Multidimensional utilities. In R. M. Thrall, C. H. Coombs, and R. L. Davis (Eds.), *Decision processes*. Wiley, New York.

Herstein, I. N. and Milnor, J. (1953). An axiomatic approach to measurable utility. *Econometrica*, **21**, 291–297.

Hölder, O. (1901). Die Axiome der Quantität und die Lehre vom Mass. *Berichte Verhand. König. Sächs. Gesell. Wiss.* (Leipzig), *Math. Phys. Cl.*, **53**, 1–64.

Hurwicz, L. and Richter, M. K. (1970). Revealed preference without demand continuity assumptions. In J. S. Chipman, L. Hurwicz, M. K. Richter, and H. Sonnenschein (Eds.), *Studies in the mathematical foundations of utility and demand theory: a symposium at the University of Michigan, 1968*. Harcourt, Brace and World, Inc.

Jensen, N. E. (1967). An introduction to Bernoullian utility theory. I. Utility functions. *Swedish Journal of Economics*, **69**, 163–183.

Kannai, Y. (1963). Existence of a utility in infinite dimensional partially ordered spaces. *Israel Journal of Mathematics*, **1**, 229–234.

Kelley, J. L. (1955). *General topology*. Van Nostrand, Princeton, New Jersey.

Koopmans, T. C. (1960). Stationary ordinal utility and impatience. *Econometrica*, **28**, 287–309.

———, Diamond, P. A. and Williamson, R. E. (1964). Stationary utility and time perspective. *Econometrica*, **32**, 82–100.

Kraft, C. H., Pratt, J. W. and Seidenberg, A. (1959). Intuitive probability on finite sets. *Annals of Mathematical Statistics*, **30**, 408–419.

Krantz, D. H. (1964). Conjoint measurement: the Luce-Tukey axiomatization and some extensions. *Journal of Mathematical Psychology*, **1**, 248–277.

——— (1967). A survey of measurement theory. *Michigan Mathematical Psychology Program*, MMPP 67-4. University of Michigan, Ann Arbor.

Kyburg, H. E. and Smokler, H. E. (Eds., 1964). *Studies in subjective probability*. Wiley, New York.

Loève, M. (1960). *Probability theory*, 2nd Ed. Van Nostrand, Princeton, New Jersey.

Luce, R. D. (1956). Semiorders and a theory of utility discrimination. *Econometrica*, **24**, 178–191.

——— (1966). Two extensions of conjoint measurement. *Journal of Mathematical Psychology*, **3**, 348–370.

——— and Raiffa, H. (1957). *Games and decisions*. Wiley, New York.

——— and Suppes, P. (1965). Preference, utility, and subjective probability. In R. D. Luce, R. R. Bush, and E. Galanter (Eds.), *Handbook of Mathematical Psychology*, III. Wiley, New York.

——— and Tukey, J. W. (1964). Simultaneous conjoint measurement: a new type of fundamental measurement. *Journal of Mathematical Psychology*, **1**, 1–27.

Marschak, J. (1950). Rational behavior, uncertain prospects, and measurable utility. *Econometrica*, **18**, 111–141.

Newman, P. and Read, R. (1961). Representation problems for preference orderings. *Journal of Economic Behavior*, **1**, 149–169.

Pareto, V. (1927). *Manuel d'economie politique*, 2nd Ed. Giard, Paris.

Pfanzagl, J. (1959). A general theory of measurement: applications to utility. *Naval Research Logistics Quarterly*, **6**, 283–294.

Pratt, J. W., Raiffa, H. and Schlaifer, R. (1964). The foundations of decision under

uncertainty: an elementary exposition. *Journal of the American Statistical Association*, **59**, 353–375.

Rader, J. T. (1963). The existence of a utility function to represent preferences. *Review of Economic Studies*, **30**, 229–232.

Raiffa, H. (1961). Risk, ambiguity, and the Savage axioms: comment. *Quarterly Journal of Economics*, **75**, 690–694.

——— and Schlaifer, R. (1961). *Applied statistical decision theory*. Division of Research, Harvard Business School, Boston.

Ramsey, F. P. (1931). Truth and probability. In F. P. Ramsey, *The foundations of mathematics and other logical essays*. Harcourt, Brace and Co., New York. Reprinted in Kyburg and Smokler (1964).

Richter, M. K. (1966). Revealed preference theory. *Econometrica*, **34**, 635–645.

Roberts, F. S. (1969). Indifference graphs. In F. Harary (Ed.), *Proof techniques in graph theory*. Academic Press, New York.

Savage, L. J. (1954). *The foundations of statistics*. Wiley, New York.

Scott, D. (1964). Measurement structures and linear inequalities. *Journal of Mathematical Psychology*, **1**, 233–247.

——— and Suppes, P. (1958). Foundational aspects of theories of measurement. *Journal of Symbolic Logic*, **23**, 113–128.

Suppes, P. (1956). The role of subjective probability and utility in decision-making. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955*, **5**, 61–73.

——— and Winet, M. (1955). An axiomatization of utility based on the notion of utility differences. *Management Science*, **1**, 259–270.

——— and Zinnes, J. L. (1963). Basic measurement theory. In R. D. Luce, R. R. Bush, and E. Galanter (Eds.), *Handbook of Mathematical Psychology*, I. Wiley, New York.

Swalm, R. O. (1966). Utility theory—insights into risk taking. *Harvard Business Review*, November–December, 123–136.

Szpilrajn, E. (1930). Sur l'extension de l'ordre partiel. *Fundamenta Mathematicae*, **16**, 386–389.

Thielman, H. P. (1953). *Theory of functions of real variables*. Prentice-Hall, Englewood Cliffs, New Jersey.

Thomsen, G. (1927). Un teorema topologico sulle schiere di curve e una caratterizzazione geometrica delle superficie isotermoasintotiche. *Bollettino della Unione Matematica Italiana*, **6**, 80–85.

Tucker, A. W. (1956). Dual systems of homogeneous linear relations. In H. W. Kuhn and A. W. Tucker (Eds.). *Linear inequalities and related systems*, *Annals of Mathematics Study 38*. Princeton University Press, Princeton, New Jersey.

Tversky, A. (1964). Finite additive structures. *Michigan Mathematical Psychology Program*, MMPP 64-6. University of Michigan, Ann Arbor.

——— (1967). A general theory of polynomial conjoint measurement. *Journal of Mathematical Psychology*, **4**, 1–20.

von Neumann, J. and Morgenstern, O. (1947). *Theory of games and economic behavior*, 2nd Ed. Princeton University Press, Princeton, New Jersey.

Weldon, J. C. (1950). A note on measures of utility. *Canadian Journal of Economics and Political Science*, **16**, 227–233.

Williams, A. C. and Nassar, J. I. (1966). Financial measurement of capital investments. *Management Science*, **12**, 851–864.

Winkler, R. L. (1967*a*). The assessment of prior distributions in Bayesian analysis. *Journal of the American Statistical Association*, **62**, 776–800.

——— (1967*b*). The quantification of judgment: some methodological suggestions. *Journal of the American Statistical Association*, **62**, 1105–1120.

Wold, H. (1943). A synthesis of pure demand analysis: I, II, III. *Skandinavisk Aktuarietidskrift*, **26**.

——— and Jureen, L. (1953). *Demand analysis*. Wiley, New York.

Yokoyama, T. (1956). Continuity conditions of preference ordering. *Osaka Economic Papers*, **4**, 39–45.

AUTHOR INDEX

SUBJECT INDEX

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
Research Analysis Corporation McLean, Virginia 22101	UNCLASSIFIED
	2b. GROUP

3. REPORT TITLE

UTILITY THEORY FOR DECISION MAKING

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Final report

5. AUTHOR(S) (First name, middle initial, last name)

Peter C. Fishburn

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
June 1970	224	107

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
DA: No. 44-188-ARO-1	RAC-R-105
b. PROJECT NO. ONR: No. N00014-67-C-0434	
c. 090.101	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY

13. ABSTRACT

The book presents a concise yet mathematically complete treatment of modern utility theories that covers nonprobabilistic preference theory, the von Neumann-Morgenstern expected-utility theory and its extensions, and the joint axiomatization of utility and subjective probability. ( )

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
decision						
utility						
value						
probability						
preference						